

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	327	100.0	482	4	US-09-513-999C-3894	Sequence 3894, Ap
2	327	100.0	507	4	US-09-949-016-4940	Sequence 4940, Ap
3	253.6	77.6	3208	4	US-09-780-016-27	Sequence 27, Appl
4	253.6	77.6	3208	4	US-10-214-811-27	Sequence 27, Appl
5	170	52.0	402	4	US-09-513-999C-10371	Sequence 10371, A
6	170	52.0	463	4	US-09-621-976-15180	Sequence 15180, A
7	153.2	46.9	411	4	US-09-640-211A-1731	Sequence 1731, Ap
8	138.2	42.3	490	4	US-09-270-767-26812	Sequence 26812, A
c	138.2	42.3	1101	4	US-09-270-767-11265	Sequence 11265, A
10	114.4	35.0	357	4	US-09-248-796A-5495	Sequence 5495, Ap
11	92.6	28.3	25274	4	US-09-949-016-16682	Sequence 16682, A
12	90	27.5	301	4	US-09-313-294A-492	Sequence 492, App
13	74	22.6	342	4	US-09-826-312A-7	Sequence 7, Appli
14	74	22.6	342	4	US-09-542-497A-7	Sequence 7, Appli
15	46.4	14.2	439	4	US-09-799-451-296	Sequence 296, App
16	36	11.0	601	4	US-09-949-016-174803	Sequence 174803, A
17	33.6	10.3	87734	4	US-09-949-016-17521	Sequence 17521, A
18	33.4	10.2	170	4	US-09-270-767-26253	Sequence 26253, A
c	33.4	10.2	170	4	US-09-270-767-10788	Sequence 10788, A
19	33.4	10.2	585	4	US-09-270-767-10788	Sequence 10788, A
20	33	10.1	601	4	US-09-949-016-120325	Sequence 120325, A
c	33	10.1	219964	4	US-09-949-016-15086	Sequence 15086, A
22	32.6	10.0	96845	4	US-09-949-016-13658	Sequence 13658, A
23	31	9.5	2408	1	US-08-608-241-1	Sequence 1, Appli
24	31	9.5	2408	2	US-08-922-182-1	Sequence 1, Appli
25	31	9.5	2408	2	US-08-919-953-1	Sequence 1, Appli
26	31	9.5	2408	3	US-09-192-983-1	Sequence 1, Appli
27	30.4	9.3	119762	4	US-09-949-016-17313	Sequence 17313, A



## US-10-214-811-27

Query Match 77.6%; Score 253.6; DB 4; Length 3208;  
Best Local Similarity 98.5%; Pred. No. 1.7e-80;  
Matches 256; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 68 AAGTGAAGAGTGGGATGAGTGGGCTGGGCTGGGATATGTTGATTAACCTGG 127  
Db 2765 AAAAAAAAAATGGAATGAGTGGGCTGGGCTGGGATATGTTGATTAACCTGG 2824  
Qy 128 CCATCTGCAGGAACACATTTATGATCTTTGATAGATGTCAGCTAACCGGCGTCCG 187  
Db 2825 CCACTCGAGNACCACATTAAGATCTTTGATAGATGTCAGCTAACCGGCGTCCG 2884  
Qy 188 CTACTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAACATGCTTTTCACTTCCACT 247  
Db 2885 CTACTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAACATGCTTTTCACTTCCACT 2944  
Qy 248 GCATCTCTGGCTGCTCAAAACACAGAGTGTGTCATTTGGACAACAGAGTGGGAAT 307  
Db 2945 GCATCTCTGGCTGCTCAAAACACAGAGTGTGTCATTTGGACAACAGAGTGGGAAT 3004  
Qy 308 TCCAAAAGTATGGGCACTAG 327  
Db 3005 TCCAAAAGTATGGGCACTAG 3024

## RESULT 5

US-09-513-999C-10371  
; Sequence 10371, Application US/09513999C  
; Patent No. 6783961

; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 10371  
; LENGTH: 402

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 20  
; OTHER INFORMATION: k=g or t  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 38  
; OTHER INFORMATION: s=g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 79  
; OTHER INFORMATION: n=a, g, c or t  
US-09-513-999C-10371

Query Match 52.0%; Score 170; DB 4; Length 402;  
Best Local Similarity 100.0%; Pred. No. 7.2e-51;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 GCATAGATGTCAAGCTAACCGGCGTCCGCTACTTTCAGAGAGTGTACTGCGATCGG 217  
Db 106 GCATAGATGTCAAGCTAACCGGCGTCCGCTACTTTCAGAGAGTGTACTGCGATCGG 165  
Qy 218 GAGTCTGTAAACATGCTTTTCACTTCCATCTCTCGCTGGCTCAAAACACGACAGG 277  
Db 166 GAGTCTGTAAACATGCTTTTCACTTCCATCTCTCGCTGGCTCAAAACACGACAGG 225

Qy 278 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCTCAAAAGTATGGGCACTAG 327  
Db 226 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCTCAAAAGTATGGGCACTAG 275

## RESULT 6

US-09-621-976-15180  
; Sequence 15180, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15180  
; LENGTH: 463

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 74  
; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-15180

Query Match 52.0%; Score 170; DB 4; Length 463;  
Best Local Similarity 100.0%; Pred. No. 7.9e-51;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 GCATAGATGTCAAGCTAACCGGCGTCCGCTACTTTCAGAGAGTGTACTGCGATCGG 217  
Db 101 GCATAGATGTCAAGCTAACCGGCGTCCGCTACTTTCAGAGAGTGTACTGCGATCGG 160  
Qy 218 GAGTCTGTAAACATGCTTTTCACTTCCATCTCTCGCTGGCTCAAAACACGACAGG 277  
Db 161 GAGTCTGTAAACATGCTTTTCACTTCCATCTCTCGCTGGCTCAAAACACGACAGG 220

Qy 278 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCTCAAAAGTATGGGCACTAG 327  
Db 221 TGTGTCCATTGGACAACAGAGAGTGGGAATTCCTCAAAAGTATGGGCACTAG 270

## RESULT 7

US-09-640-211A-1731  
; Sequence 1731, Application US/09640211A  
; Patent No. 6833446

; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021C1U  
; CURRENT APPLICATION NUMBER: US/09/640,211A  
; CURRENT FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 2388  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1731  
; LENGTH: 411

; TYPE: DNA  
; ORGANISM: Pinus radiata  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(411)  
; OTHER INFORMATION: n = A, T, C or G  
US-09-640-211A-1731





Db 269 ATTACATTGATTAGAGATGGTTGAAACAAAGAAATGCATGTCCTTGGATAGTACTA 328  
 Qy 299 AGTGGGAATTCAAAAGATGAGCACTA 326  
 Db 329 ATTGGACTTATCAAAAATTTGGGTAATTA 356

RESULT 11  
 US-09-949-016-16682  
 ; Sequence 16682, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 16682  
 ; LENGTH: 25274  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-16682

Query Match 28.3%; Score 92.6; DB 4; Length 25274;  
 Best Local Similarity 96.0%; Pred. No. 6.7e-22;  
 Matches 95; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 221 TCTGTAAACATGCTTTTCACTTCCACTGTCATCTCTCGCTGGCTCAAAAACACGACAGGTGT 280  
 Db 18403 TCTTTGACATGCTTTTCACTTCCACTGTCATCTCTCGCTGGCTCAAAAACACGACAGGTGT 18462

Qy 281 GTCCATTGACACAGAGAGTGGGAATTCCAAAGTATG 319  
 Db 18463 GTCCATTGACACAGAGAGTGGGAATTCCAAAGTATG 18501

RESULT 12  
 US-09-313-294A-492  
 ; Sequence 492, Application US/09313294A  
 ; Patent No. 6476212  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lalgudi, Raghunath V.  
 ; APPLICANT: Ito, Laura Y.  
 ; APPLICANT: Sherman, Bradley K.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
 ; FILE REFERENCE: PL-0017 US  
 ; CURRENT APPLICATION NUMBER: US/09/313,294A  
 ; CURRENT FILING DATE: 1999-05-14  
 ; NUMBER OF SEQ ID NOS: 7600  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 492  
 ; LENGTH: 301  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No. 6476212 700549333H1  
 US-09-313-294A-492

Query Match 27.5%; Score 90; DB 4; Length 301;  
 Best Local Similarity 74.0%; Pred. No. 4.4e-22;  
 Matches 114; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 38 CCAACAGCGCGCGGGAAGAGCGCTTTGAAGTCAAAAAGTGAATGCAGTACGCTCT 97  
 Db 146 CCTCTCCCGCAAGCCCAAGAGCGCTTCGAGATCAAGAGTGAACCGCTCGGCTCT 205

Qy 98 GGGCTGGGATATTGTGTTGATTAACATGTGCGCATCTGCAGGAACCAATATGATCTTT 157  
 Db 206 GGGCATGGGATATGTCGTCGCAACATGCGCTATCTGCCGCAACCAATCATGATCTAT 265

Qy 158 GCATAGAAATGTCAGCTTAACCAAGCGGTCCGCTAC 191  
 Db 266 GCATCGAGTGCAGCGCAACCAAGCCAGCGCGAC 299

RESULT 13  
 US-09-826-312A-7  
 ; Sequence 7, Application US/09826312A  
 ; Patent No. 6737244  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Issakani, Sarkiz D.  
 ; APPLICANT: Huang, Jianing  
 ; APPLICANT: Sheung, Julie  
 ; APPLICANT: Pray, Todd R.  
 ; APPLICANT: Rigel Pharmaceuticals, Inc.  
 ; TITLE OF INVENTION: Ubiquitin Ligase Assay  
 ; FILE REFERENCE: 021044-007010US  
 ; CURRENT APPLICATION NUMBER: US/09/826,312A  
 ; CURRENT FILING DATE: 2001-04-03  
 ; PRIOR APPLICATION NUMBER: US 09/542,497  
 ; PRIOR FILING DATE: 2000-04-03  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 342  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: RING finger protein ROC2  
 US-09-826-312A-7

Query Match 22.6%; Score 74; DB 4; Length 342;  
 Best Local Similarity 58.4%; Pred. No. 2.8e-16;  
 Matches 153; Conservative 0; Mismatches 100; Indels 9; Gaps 1;

Qy 52 GCGAAGAGCGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 111  
 Db 76 GCGACAGAGATGTTCTCCCTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 135

Qy 112 GTGGTTGATRACTGTCCATCTGCAGGAACCAATATGATCTTTGCATAGAAATGTCAA 171  
 Db 136 GAGTGGGATACGTGGCCCATCTGCAGGGTCCAGGTGATGGATGCTTCTTAGATGTCAA 195

Qy 172 GCTAACACGCGCTCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCAT 231  
 Db 196 GCTGAAA-----CAACAAGAGGACTGTGTGTGGTCTGGGAGAAATGTAATCAT 246

Qy 232 GCTTTTCATTCCTCACTGATCTCTCGTGGTCTCAAAAACACAGAGGTGTGTCATTTGAC 291  
 Db 247 TCCITTCACAACTGTCATCTCCCTGTGGTGAACAGAAACAAATCGCTGCTCTCTGC 306

Qy 292 AACAGAGTGGGAATTCCTAAA 313  
 Db 307 CAGCAGGACTGGGTGGTCCAAA 328

RESULT 14  
 US-09-542-497A-7  
 ; Sequence 7, Application US/09542497A  
 ; Patent No. 6740495  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Issakani, Sarkiz D.  
 ; APPLICANT: Huang, Jianing  
 ; APPLICANT: Sheung, Julie  
 ; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 24, 2005, 08:28:03 ; Search time 156 Seconds  
(without alignments)  
1132.807 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVTPSGTNSGACKK.....KTRQVCPLDNREWFQKYGH 108

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09541462/runat\_23062005\_122624\_9498/app\_query.fasta\_1.263  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DLOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09541462 @CEN 1 1 69 @runat\_23062005\_122624\_9498 -NCPV=6 -ICPV=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	482	US-09-513-999C-3894	Sequence 3894, Ap
2	616	100.0	507	US-09-949-016-4940	Sequence 4940, Ap
3	501.5	81.4	3208	US-09-780-016-27	Sequence 27, Appl
4	501.5	81.4	3208	US-10-214-811-27	Sequence 27, Appl
5	479	77.8	411	US-09-640-211A-1731	Sequence 1731, Ap
6	400.5	65.0	490	US-09-270-767-26812	Sequence 26812, A
7	400.5	65.0	1101	US-09-270-767-11265	Sequence 11265, A
8	394	64.0	402	US-09-513-999C-10371	Sequence 10371, A
9	394	64.0	463	US-09-621-976-15180	Sequence 15180, A
10	375.5	61.0	357	US-09-248-796A-5495	Sequence 5495, Ap
11	287	46.6	342	US-09-826-312A-7	Sequence 7, Appl
12	287	46.6	342	US-09-542-497A-7	Sequence 7, Appl

13	262.5	42.6	301	4	US-09-313-294A-492	Sequence 492, App
14	217	35.2	648	4	US-09-599-360B-27	Sequence 27, Appl
15	200.5	32.5	534	4	US-09-621-976-1817	Sequence 1817, Ap
16	198	32.1	671	4	US-09-621-976-1854	Sequence 1854, Ap
17	197.5	32.1	539	4	US-09-621-976-2051	Sequence 2051, Ap
18	191	31.0	654	4	US-09-621-976-1945	Sequence 1945, Ap
19	175	28.4	25274	4	US-09-949-016-16682	Sequence 16682, A
20	173	28.1	585	4	US-09-270-767-10788	Sequence 10788, A
21	150.5	24.4	439	4	US-09-799-451-296	Sequence 296, App
22	146	23.7	170	4	US-09-270-767-26253	Sequence 26253, A
23	91	14.8	940	4	US-09-023-655-667	Sequence 667, App
24	91	14.8	1839	4	US-09-828-303-10	Sequence 10, Appl
25	90	14.6	872	4	US-09-774-528-304	Sequence 304, App
26	90	14.6	893	4	US-09-949-016-4980	Sequence 4980, Ap
27	90	14.6	1690	4	US-09-828-303-2	Sequence 2, Appl1
28	90	14.6	3140	4	US-09-774-528-255	Sequence 255, App
29	90	14.6	8438	1	US-07-945-283-1	Sequence 1, Appl1
30	88	14.3	1183	4	US-09-799-451-763	Sequence 763, App
31	86	14.0	1267	4	US-09-640-211A-1319	Sequence 1319, Ap
32	86	14.0	1267	4	US-09-949-016-378	Sequence 378, App
33	86	14.0	1267	4	US-09-949-016-2914	Sequence 2914, Ap
34	85.5	13.9	3304	4	US-09-799-451-220	Sequence 220, App
35	85	13.8	1621	4	US-09-023-655-20	Sequence 20, Appl
36	85	13.8	1995	4	US-09-949-016-3134	Sequence 3134, Ap
37	85	13.8	2339	3	US-09-268-140-11	Sequence 11, Appl
38	85	13.8	2505	3	US-09-268-140-1	Sequence 1, Appl1
39	85	13.8	2517	3	US-09-268-140-7	Sequence 7, Appl1
40	85	13.8	16573	4	US-09-949-016-14876	Sequence 14876, A
41	84	13.6	3260	4	US-09-270-767-10326	Sequence 10326, A
42	83.5	13.6	624	4	US-09-270-767-963	Sequence 963, App
43	83.5	13.6	624	4	US-09-270-767-16245	Sequence 16245, A
44	83	13.5	315	3	US-09-325-932A-4	Sequence 4, Appl1
45	81.5	13.2	4259	2	US-08-816-155B-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1  
US-09-513-999C-3894  
; Sequence 3894, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59 US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3894  
; LENGTH: 482  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 29..352  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 401  
; OTHER INFORMATION: r-a or g  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 404  
; OTHER INFORMATION: m-a or c  
US-09-513-999C-3894  
Alignment Scores: 1.7e-69 Length: 482  
Pred. No.: 482

Db 67 CGCTTTGAAGTGAAGAAGTGAATGACAGTAGCCCTCTGGCCCTGGGATATTGTGTTGAT 126  
 QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
 Db 127 AACTGTGCCATCTGCAGGAACACCAATTATGGATCTTTGCATAGAAATGCTAAGCTAACCG 186  
 QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
 Db 187 GCGTCGCGTACTTTCAGAGAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 246  
 QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysPro-euAspAsnAtrGlu 100  
 Db 247 TTCACATGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCATTTGGACAAACAGAG 306  
 QY 101 TrpGluPheGlnLysTrpGlyHis 108  
 Db 307 TGGGAATTCAAAAGTATGGGCAC 330  
 RESULT 3  
 US-09-780-016-27  
 ; Sequence 27, Application US/097800016  
 ; Patent No. 6509456  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Donoho, Gregory  
 ; APPLICANT: Scoville, John  
 ; APPLICANT: Turner, C. Alexander Jr.  
 ; APPLICANT: Friedrich, Glenn  
 ; APPLICANT: Abuin, Alejandro  
 ; APPLICANT: Zambrowicz, Brian  
 ; APPLICANT: Sands, Arthur I.  
 ; TITLE OF INVENTION: No. 6509456el Human Proteases and  
 ; TITLE OF INVENTION: Polynucleotides Encoding the Same  
 ; FILE REFERENCE: LEX-0132-USA  
 ; CURRENT APPLICATION NUMBER: US/09/780,016  
 ; CURRENT FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: US 60/181,294  
 ; PRIOR FILING DATE: 2000-02-11  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Fast-Seq for Windows Version 4.0  
 ; SEQ ID NO 27  
 ; LENGTH: 3208  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-09-780-016-27  
 Alignment Scores:  
 Pred. No.: 1.03e-53 Length: 3208  
 Score: 501.50 Matches: 88  
 Percent Similarity: 95.70% Conservative: 1  
 Best Local Similarity: 94.62% Mismatches: 1  
 Query Match: 81.41% Indels: 3  
 DB: 4 Gaps: 1  
 US-09-541-462B-2 (1-108) x US-09-780-016-27 (1-3208)  
 QY 16 GlyAlaGlyLysLysArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrp 35  
 Db 2752 GGAGCATAAAAA-----AAAAATGGAATGAGTAGCCCTCTGGCCCTGG 2802  
 QY 36 AspIleValValAspAsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGlu 55  
 Db 2803 GATATTGTGGTTGATAACTGTGCGCATCTGCAGGAACACACATTTATGGATCTTTGCATAGAA 2862  
 QY 56 CysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCys 75  
 Db 2863 TGTCACAGCTAACGAGCGCTCCGTACTTTCAGAGAGTGTAATGTCGCGATGGGGAGTCTGT 2922  
 QY 76 AsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThrArgGlnValCysPro 95  
 Db 2923 AACCATGCTTTTCACTTCCACTGCACTCTCTGCTGGCTCAAAACACGACAGGTGTGTCCA 2982  
 QY 96 LeuAspAsnArgGluTrpGluPheGlnLysTrpGlyHis 108

```

; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Pinus radiata
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(411)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1731

Alignment Scores:
Pred. No.: 4,28e-52 Length: 411
Score: 479.00 Matches: 90
Percent Similarity: 84.96% Conservative: 6
Best Local Similarity: 79.65% Mismatches: 4
Query Match: 77.76% Indels: 13
DB: 4 Gaps: 3

US-09-541-462B-2 (1-108) x US-09-640-211A-1731 (1-411)
QY 3 AlaAlaMetAspValAsp-----ThrProSer-----GlyThrIenSerGlyAla 17
Db 74 GCTTCAACAGACATAGATATGATTCGCGTGCCTCTGGCGAGGGTTCCAGCTCTCAAGCG 133
QY 18 Gly-----LysLysArgPheGluValLysLysValThrAsnAlaVal 30
Db 134 GGACCAACGGCTTCCACCAGAAGCCAAACGTTTCGAAATCAAGAAGTGAAGTAAATGCTGTA 193
QY 31 AlaLeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIleMet 50
Db 194 GCCTTTGGCGTGGGATATTGTGTGATTAATTTGTCATTTTCAGAAACCATCATG 253
QY 51 AspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrVal 70
Db 254 GACCTCTGTATTAGTGTGAGTGTGAGCAAAATCAAGCAAGTGCACAACAAAGTGAAGTAAATGCTACTGT 313
QY 71 AlaTrpGlyValCysAsnHisAlaPheHisCysIleSerArgTrpLeuLysThr 90
Db 314 GCATGGGGTGTTCGAATCAGCGCTTTTCATTTTCATTGTCATAAGTCCGTGGCTCAAGACA 373
QY 91 ArgGlnValCysProLeuAspAsnArg-GluTrpGlu 102
Db 374 CGCAACAGTCTGCCCATTAGATNATAAGTAGTGGAG 410

RESULT 6
US-09-270-767-26812/c
; Sequence 26812, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26812
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-26812

Alignment Scores:
Pred. No.: 5,82e-42 Length: 490
Score: 400.50 Matches: 70
Percent Similarity: 77.57% Conservative: 13
Best Local Similarity: 65.42% Mismatches: 19
Query Match: 65.02% Indels: 5
DB: 4 Gaps: 2

```

DD 201 1C1CG1GG1CAHFF  
Q: 105 100T0001WH18 108

```

; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5495
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-5495

Alignment Scores:
Pred. No.: 5,82e-39 Length: 357
Score: 375.50 MatchChes: 62
Percent Similarity: 73.27% Conservative: 12
Best Local Similarity: 61.39% MismatchChes: 26
Query Match: 60.96% Indels: 1
DB: 4 Gaps: 1

US-09-541-462B-2 (1-108) x US-09-248-796A-5495 (1-357)

Qy 9 ThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPheGluValLysLysTyrAsn 28
Db 52 ACAACAGAACCAACATCAAAACCACATCAAAACCAAGATTTGAAGTGAAGAAATGGACA 111
Qy 29 AlaValAlaLeuTTPAlaTTPAspIleValValAspAsnCysAlaIleCysArgAsnHis 48
Db 112 GCGGTAGCGCTTTTGGTCTATGGGATATGCAAATTGAAAATTGTGCCATTTGTAGAAATCAT 171
Qy 49 IleMetAspLeuCysIleGluCysGlnAlaAsn--GlnAlaSerAlaThrSerGluGlu 67
Db 172 TTAATGGACCAACGATTAATGAATGTCACCAATGCTATGGGTATATATTCCTTCAGAAGAG 231
Qy 68 CysThrValAlaTTPGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTTP 87
Db 232 TGTATTCTGCTTGGGGAGTATGTAATCATGCAATTCATTTACATTTGATTAGAAGATGG 291
Qy 88 LeuLysThrArgGlnValCysProLeuAspAsnArgGluTTPGluPheGlnLysTyrGly 107
Db 292 TTCAAAAACAAGAAATGCAATGTCCTTGGTAGTACTAATTGGACTTATCAAAAAATGGGT 351
Qy 108 His 108
Db 352 AAT 354

RESULT 11
US-09-826-312A-7
; Sequence 7, Application US/09826312A
; Patent No. 6737244
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/09/826,312A
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: RING finger protein ROC2
US-09-826-312A-7

Alignment Scores:
Pred. No.: 1.1e-27 Length: 342
Score: 287.00 MatchChes: 48
Percent Similarity: 63.92% Conservative: 14

```

Best Local Similarity: 49.48% Mismatches: 31  
Query Match: 46.59% Indels: 4  
DB: 4 Gaps: 2

US-09-541-462B-2 (1-108) x US-09-826-312A-7 (1-342)

QY 11 SerGlyThrAsnSerGlyAlaGlyLysLysAspPheGluValLysLysTrpAsnAlaVal 30  
DB 58 TCAGGCTCAACGTCGGA---GGCGACAGATGTTCTCCCTCAAGAAGTGGAAACCCGGTG 114  
QY 31 AlaLeuTrpAlaTrpAspLeuValValAspAsnCysAlaIleCysArgAsnHisIleMet 50  
DB 115 GCCATGTGGAGCTGGGACGTGGAGTACGTGGCCATCTGCAGGGTCCAGGTGATG 174  
QY 51 AspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrVal 70  
DB 175 GATGCTGCTTGTAGATGTCAAGCTGAAAC-----AAACAAGAGGACTGTGTTGTG 225  
QY 71 AlaTrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThr 90  
DB 226 GTCTGGGAGATGTAATCATCTCTCCACAACCTGCTGCATGCTCCCTGTGGTGAACAG 285  
QY 91 ArgGlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGly 107  
DB 286 AACAAATCGCTGCCCTCTCTGCCAGCAGGACTGGGTGTCACAAAGAATCGGC 336

RESULT 12

US-09-542-497A-7

; Sequence 7, Application US/09542497A  
; Patent No. 6740495  
; GENERAL INFORMATION:  
; APPLICANT: Isaakani, Sarkiz D.  
; APPLICANT: Huang, Jianing  
; APPLICANT: Shuang, Julie  
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY  
; FILE REFERENCE: A-68613/RMS/JJD  
; CURRENT APPLICATION NUMBER: US/09/542,497A  
; CURRENT FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 342  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-542-497A-7

Alignment Scores:  
Pred. No.: 1.1e-27 Length: 342  
Score: 287.00 Matches: 48  
Percent Similarity: 63.92% Conservative: 14  
Best Local Similarity: 49.48% Mismatches: 31  
Query Match: 46.59% Indels: 4  
DB: 4 Gaps: 2

US-09-541-462B-2 (1-108) x US-09-542-497A-7 (1-342)

QY 11 SerGlyThrAsnSerGlyAlaGlyLysLysAspPheGluValLysLysTrpAsnAlaVal 30  
DB 58 TCAGGCTCAACGTCGGA---GGCGACAGATGTTCTCCCTCAAGAAGTGGAAACCCGGTG 114  
QY 31 AlaLeuTrpAlaTrpAspLeuValValAspAsnCysAlaIleCysArgAsnHisIleMet 50  
DB 115 GCCATGTGGAGCTGGGACGTGGAGTACGTGGCCATCTGCAGGGTCCAGGTGATG 174  
QY 51 AspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrVal 70  
DB 175 GATGCTGCTTGTAGATGTCAAGCTGAAAC-----AAACAAGAGGACTGTGTTGTG 225  
QY 71 AlaTrpGlyValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThr 90  
DB 226 GTCTGGGAGATGTAATCATCTCTCCACAACCTGCTGCATGCTCCCTGTGGTGAACAG 285  
QY 91 ArgGlnValCysProLeuAspAsnArgGluTrpGluPheGlnLysTyrGly 107

DB 286 AACAAATCGCTGCCCTCTCTGCCAGCAGGACTGGGTGTCACAAAGAATCGGC 336

RESULT 13

US-09-313-294A-492

; Sequence 492, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 492  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700549333H1  
US-09-313-294A-492

Alignment Scores:  
Pred. No.: 1.24e-24 Length: 301  
Score: 262.50 Matches: 51  
Percent Similarity: 71.43% Conservative: 4  
Best Local Similarity: 66.23% Mismatches: 7  
Query Match: 42.61% Indels: 15  
DB: 4 Gaps: 2

US-09-541-462B-2 (1-108) x US-09-313-294A-492 (1-301)

QY 3 AlaAlaMetAspValAsp-----ThrProSerGlyThrAsnSer 15  
DB 70 TCCGCCATGGAGACCGACATCAAGCGCGCGCGCCGCCAGCTGGCGAGGATCC 129  
QY 16 GlyAlaGly-----LysLysArgPheGluValLysLysTrp 27  
DB 130 TCTGCGCGCGTCCGTCCTCTCCCAAGCCCAACAAAGCGCTTCGAGATCAAGAAGTGG 189  
QY 28 AsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAlaIleCysArgAsn 47  
DB 190 AACGCCGTGCGCTCTGGCGATGGGATATCGTCTCGACAACTGCGCTATCTGCCGCAAC 249  
QY 48 HisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAlaThr 64  
DB 250 CACATCATGATCTATGATCGAGTCCAGGCGCAACCAAGCCAGCGGACC 300

RESULT 14

US-09-599-360B-27

; Sequence 27, Application US/09599360B  
; Patent No. 6548633  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Bougueleret, L.  
; APPLICANT: Jobert, S.  
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides  
; FILE REFERENCE: GENSET.050CP3  
; CURRENT APPLICATION NUMBER: US/09/599,360B  
; CURRENT FILING DATE: 2000-06-21  
; PRIOR APPLICATION NUMBER: 60/113,686  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: 60/141,032  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 09/469,099  
; PRIOR FILING DATE: 1999-12-21  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: Patent.pm  
; SEQ ID NO 27



```
; LENGTH: 648
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 187..438
; NAME/KEY: polyA signal
; LOCATION: 612..617
; NAME/KEY: polyA site
; LOCATION: 632..648
US-09-599-360B-27
Alignment Scores:
Pred. No.: 2,368-18 Length: 648
Score: 217.00 Matches: 43
Percent Similarity: 44.53% Conservative: 18
Best Local Similarity: 31.39% Mismatches: 40
Query Match: 35.23% Indels: 36
DB: 4 Gaps: 3
US-09-541-462B-2 (1-108) x US-09-599-360B-27 (1-648)
Qy 2 AlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValAlaGlyLysLysArg 21
Db 49 GCCGCGACTGTGGTCTTTTATACCTTCCCGCGGACGCGCGCTGCCAACGGAAGG 108
Qy 22 PheGlu-----PheGluValLysLysTrpAsnAlaValAlaLeuTrpAla 34
Db 109 GCGGAGACGGAGTTTCGTATGTTGCCAGGCCCATTTGAGATCTTTGAAGATATCCTCA 168
Qy 24 -----ValLysLysTrpAsnAlaValAlaLeuTrp 33
Db 169 ACGTGAGGCTCTGTGCCATGAGTGAGTAAAGTAAAGTCTGGAACGGCGTGCCACTTGG 228
Qy 34 AlaTrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIleMetAspLeuCys 53
Db 229 CTCCTGGTGGCCACAGATGAGACTGTGGCATCTGCAGGATGGCATTTAACGGATGCTGC 288
Qy 54 IleGluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAlaTrpGly 73
Db 289 CTGACTGCAAG-----GTGCCCGGCGACGACTGCCGCTGTGTGGGGC 333
Qy 74 ValCysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThrArgGlnVal 93
Db 334 CAGTGTCTCCCACTGCTTCCATGATTCATTCCTCAAGTGGCTGCACGACGAGGTG 393
Qy 94 -----CysProLeuAspAsnArgGluTrpGluPheGlnLysTyGly 107
Db 394 CAGCAGCACTGCCCATGTGCCCGCAGGAGTGAAGTTCAAGAGTGAGGC 444
RESULT 15
US-09-621-976-1817
; Sequence 1817, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1817
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 106..441
; NAME/KEY: sig_peptide
; LOCATION: 106..423
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.40000009536743
; OTHER INFORMATION: seq GILILLPHAGADG/CW
US-09-621-976-1817
Alignment Scores:
Pred. No.: 2,368-16 Length: 534
Score: 200.50 Matches: 42
Percent Similarity: 51.72% Conservative: 18
Best Local Similarity: 36.21% Mismatches: 41
Query Match: 32.55% Indels: 16
DB: 4 Gaps: 3
US-09-541-462B-2 (1-108) x US-09-621-976-1817 (1-534)
Qy 2 AlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArg 21
Db 31 GCCGCGACTGTGGTCTTTTATACCTTCCCGCGGACGCGCGCTGCCAACGGAAGG 90
Qy 22 -----PheGluValLysLysTrpAsnAlaValAlaLeuTrpAla 34
Db 91 GCGGCTCTGTCTGCCATGAAGTGAAGTAAAGTCTGGAACGGCGTCCCACTTGGCTC 150
Qy 35 TrpAspIleValValAspAsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIle 54
Db 151 TGGTGGCCACGATGAGACTGTGGCATCTGCAGGATGGCATTTAACGGATGCTGCCCT 210
Qy 55 GluCysGlnAlaAsnGlnAlaSerAlaThrSerGluGluCysThrValAlaTrpGlyVal 74
Db 211 GACTGCAAG-----GTGCCCGGCGACGACTGCCGCTGTGTGGGGCCAG 255
Qy 75 CysAsnHisAlaPheHisPheHisCysIleSerArgTrpLeuLysThrArgGlnVal---- 93
Db 256 TGCTCCCACTGCTTCCACATGTCATTCCTCAAGTGGCTGCACGACGACGAGGTGCAG 315
Qy 94 -----CysProLeuAspAsnArgGluTrpGluPheGlnLysTyGly 107
Db 316 CAGCACTGCCCATGTGCCGCCA-GAATGGAAGTTCAAGAGTGAGGC 362
Search completed: June 24, 2005, 12:05:35
Job time : 159 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 15:20:31 ; Search time 43 Seconds  
(without alignments)  
187.491 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVDTSGTSGAGKK.....KTRQVCPLDNREWEFKYGH 108

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	108	4	US-09-826-312A-6
2	616	100.0	108	4	US-09-542-497A-6
3	616	100.0	108	4	US-09-513-999C-7971
4	616	100.0	110	4	US-09-949-016-10811
5	375.5	61.0	118	4	US-09-248-796A-19598
6	287	46.6	113	4	US-09-826-312A-8
7	287	46.6	113	4	US-09-542-497A-8
8	208	33.8	84	4	US-09-599-360B-77
9	202	32.8	84	4	US-09-826-312A-5
10	202	32.8	84	4	US-09-542-497A-5
11	185	30.0	112	4	US-09-621-976-5677
12	185	30.0	112	4	US-09-621-976-5805
13	105.5	17.1	94	4	US-09-621-976-5714
14	90	14.6	195	4	US-09-949-016-10851
15	86	14.0	326	4	US-09-949-016-6249
16	86	14.0	336	4	US-09-949-016-8785
17	85	13.8	664	3	US-09-268-140-2
18	85	13.8	664	4	US-09-949-016-9005
19	84	13.6	556	4	US-09-270-767-1922
20	83.5	13.6	181	4	US-09-270-767-32592
21	83.5	13.6	181	4	US-09-270-767-47809
22	83	13.5	104	4	US-09-325-932A-49
23	83	13.5	337	4	US-09-828-303-18
24	81	13.1	484	4	US-09-248-796A-15483
25	80.5	13.1	895	4	US-09-270-767-42746
26	80	13.0	160	4	US-09-949-016-8123
27	80	13.0	180	2	US-08-786-606-3

28	80	13.0	180	2	US-08-933-750C-48	Sequence 48, Appl
29	80	13.0	180	3	US-09-234-613-48	Sequence 48, Appl
30	80	13.0	180	4	US-09-949-016-11029	Sequence 11029, A
31	80	13.0	252	4	US-09-270-767-45203	Sequence 45203, A
32	80	13.0	284	2	US-08-786-606-9	Sequence 9, Appl
33	79	12.8	50	3	US-09-052-089A-15	Sequence 15, Appl
34	79	12.8	114	4	US-09-270-767-57487	Sequence 57487, A
35	79	12.8	352	4	US-09-270-767-42211	Sequence 42211, A
36	79	12.8	359	4	US-09-663-600A-106	Sequence 106, App
37	79	12.8	381	2	US-08-867-057-1	Sequence 1, Appl
38	79	12.8	381	2	US-08-867-057-3	Sequence 3, Appl
39	79	12.8	381	2	US-09-128-369-1	Sequence 1, Appl
40	79	12.8	381	2	US-09-128-369-3	Sequence 3, Appl
41	79	12.8	381	4	US-09-663-600A-200	Sequence 200, App
42	79	12.8	410	1	US-07-945-283-4	Sequence 4, Appl
43	78	12.7	448	4	US-09-949-016-9938	Sequence 9938, Ap
44	77.5	12.6	342	4	US-09-270-767-45316	Sequence 45316, A
45	77	12.5	258	4	US-09-949-016-9556	Sequence 9556, Ap

ALIGNMENTS

RESULT 1  
US-09-826-312A-6  
; Sequence 6, Application US/09826312A  
; Patent No. 6737244  
; GENERAL INFORMATION:  
; APPLICANT: Issakani, Sarkiz D.  
; APPLICANT: Huang, Jianing  
; APPLICANT: Sheung, Julie  
; APPLICANT: Pray, Todd R.  
; TITLE OF INVENTION: Ubiquitin Ligase Assay  
; FILE REFERENCE: 021044-007010US  
; CURRENT APPLICATION NUMBER: US/09/826,312A  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 09/542,497  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: RING finger protein ROC1  
US-09-826-312A-6  
Query Match 100.0%; Score 616; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 4.7e-60;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAAMDVDTSGTSGAGKKRFEVKKNVALWAWD1VVDNCAICRHHIMDLCEQANQ 60  
Db 1 MAAMDVDTSGTSGAGKKRFEVKKNVALWAWD1VVDNCAICRHHIMDLCEQANQ 60  
Qy 61 ASATSECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFKYGH 108  
Db 61 ASATSECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFKYGH 108  
RESULT 2  
US-09-542-497A-6  
; Sequence 6, Application US/09542497A  
; Patent No. 6740495  
; GENERAL INFORMATION:  
; APPLICANT: Issakani, Sarkiz D.  
; APPLICANT: Huang, Jianing  
; APPLICANT: Sheung, Julie  
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY  
; FILE REFERENCE: A-68613/RMS/JJD  
; CURRENT APPLICATION NUMBER: US/09/542,497A

```

; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-497A-6

```

	Query Match	100.0%;	Score 616;	DB 4;	Length 108;
	Best Local Similarity	100.0%;	Pred. NO. 4.7e-60;		
	Matches 108; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAMDVDTPSGTNSGAGKGRFEVKKNVALWAWDIWDNCATCRNHIMDLCTECCAAQ	60		
Dd	1	MAAAMDVDTPSGTNSGAGKGRFEVKKNVALWAWDIWDNCATCRNHIMDLCTECCAAQ	60		
Qy	61	ASATSEECTVAMGVGCNHAFFHCISRWLKTQVCPLDNREWEFOKYGH	108		
Dd	61	ASATSEECTVAMGVGCNHAFFHCISRWLKTQVCPLDNREWEFOKYGH	108		

```

RESULT 3
US-09-513-999C-7971
; Sequence 7971, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7971
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7971

```

	Query Match	100.0%;	Score 616;	DB 4;	Length 108;
	Best Local Similarity	100.0%;	Pred. No. 4.7e-60;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAMDVDTSGTNSGAGKKRFEVKKNVALWAMDIVDNCATCRNHIMDLCTECQANQ	60		
Db	1	MAAAMDVDTSGTNSGAGKKRFEVKKNVALWAMDIVDNCATCRNHIMDLCTECQANQ	60		
Qy	61	ASATSECTVAMGVCNHAHFHICISRWLKTROVCPDLNREWFQKYGH	108		
Db	61	ASATSECTVAMGVCNHAHFHICISRWLKTROVCPDLNREWFQKYGH	108		

```

RESULT 4
US-09-949-016-10811
; Sequence 10811, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASSEQ for Windows Version 4.0
; SEQ ID NO 10811
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10811

```

	Query Match	100.0%;	Score 616;	DB 4;	Length 110;
	Best Local Similarity	100.0%;	Pred. No. 4.8e-60;		
	Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAAMDVDT	PSGTSGAGKGR	FEVKKNAVAL	WAWDIVDNCALCRNHIMDLCTECQANQ 60
Db	3	MAAAMDVDT	PSGTSGAGKGR	FEVKKNAVAL	WAWDIVDNCALCRNHIMDLCTECQANQ 62
Qy	61	ASATSECTV	AGVNCNHA	PHFHCISRLW	KTRQVCPDLNREWEFKYGH 108
Db	63	ASATSECTV	AGVNCNHA	PHFHCISRLW	KTRQVCPDLNREWEFKYGH 110

```

RESULT 5
US-09-248-796A-19598
; Sequence 19598, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AM
; TITLE OF INVENTION: FOR DIAGNOSTICS AN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,7
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19598
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19598

```

Query Match	61.0%	Score	375.5;	DB	4;	Length	118;	
Best Local Similarity	61.4%;	Pred. No.	1.1e-33;					
Matches	69;	Conservative	12;	Mismatches	26;	Indels	1; Gaps	1;
Qy	9	TPSGTSGACKKFEVKKNNVALWAWD	IVDNCAICRNHIMDLCTECQAN-QASATSEE	67				
	:	:	:	:	:	:	:	
	:	:	:	:	:	:	:	
Dd	18	TTEPTSKSPKRFEVKWTAVAFWSMDQIENCAICRNHLMPECTEQPNMGNI	PSEE	77				
	:	:	:	:	:	:	:	
	:	:	:	:	:	:	:	
Qy	68	CTVAWGVCNHAFHFHCISRWLKTQRQCPLDNREWEFOKYGH	108					
	:	:	:	:	:	:	:	
	:	:	:	:	:	:	:	
Dd	78	CIPAWGVCNHAFHLCHCIRRLKTRNACPLDSTNTWTYOKLGN	118					
	:	:	:	:	:	:	:	
	:	:	:	:	:	:	:	

```

RESULT 6
US-09-826-312A-8
; Sequence 8, Application US/09826312A
; Patent No. 6737244
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/09/826,312A
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542,497

```

```

; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: RING finger protein ROC2
US-09-826-312A-8

Query Match
Best Local Similarity 46.6%; Score 287; DB 4; Length 113;
Matches 48; Conservative 14; Mismatches 31; Indels 4; Gaps 2;

QY 11 SGTNSGAGKRFVKKNAVALWAWDIVDNCALICRNHIMDLCTECQANQASATSECTV 70
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 SGSTSG-GDKWFSLLKKNPVMWSWDECDTCAICRVQVMDACLRCQAE--KQEDCV 75

QY 71 AWGVCHAFHFCISRWLKTQVCPDLNREWEFKYQ 107
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 VMGECHNSPHNCMSLWVKQNNRCPCLCQDWMVVRIG 112

RESULT 7
US-09-542-497A-8
; Sequence 8, Application US/09542497A
; Patent No. 6740495
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Shuang, Julie
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY
; FILE REFERENCE: A-68613/RMS/JJD
; CURRENT APPLICATION NUMBER: US/09/542,497A
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 8
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-497A-8

Query Match
Best Local Similarity 46.6%; Score 287; DB 4; Length 113;
Matches 48; Conservative 14; Mismatches 31; Indels 4; Gaps 2;

QY 11 SGTNSGAGKRFVKKNAVALWAWDIVDNCALICRNHIMDLCTECQANQASATSECTV 70
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 20 SGSTSG-GDKWFSLLKKNPVMWSWDECDTCAICRVQVMDACLRCQAE--KQEDCV 75

QY 71 AWGVCHAFHFCISRWLKTQVCPDLNREWEFKYQ 107
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 VMGECHNSPHNCMSLWVKQNNRCPCLCQDWMVVRIG 112

RESULT 8
US-09-599-360B-77
; Sequence 77, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bougueret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,586
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25

```

```

; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-497A-5

```

Query Match	32.8%;	Score 202;	DB 4;	Length 84;
Best Local Similarity	36.4%;	Pred. No. 7.4e-15;		
Matches	32;	Conservative 17;	Mismatches 31;	Indels 8; Gaps 2;
Qy	21	RFEVKKNAVALWAWD1VWNCIAICRNHIMDLCTECQANQASATSECTVANGVCNIAFH	80	
Db	2	KVKIKCWGVATLWLVANDENCIGICRMFNGCCPDCK----	VPGDCCPLVWGQSCSHCF	56
Qy	81	FHCISRWLKTROV----	CPLONREWEFOK	105
Db	57	MHCILKWLHACVOOHCPCMCROWTKFKE		84

```

RESULT 11
US-09-621-976-5677
; Sequence 5677, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5677
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -106...-1
US-09-621-976-5677

```

	Query Match	30.0%; Score 185; DB 4; Length 112;
	Best Local Similarity	39.0%; Pred.No.7.4e-13;
Matches	30; Conservative	13; Mismatches 28; Indels 8; Gaps 2;
Qy	21 RFEVKNNVALWALWDIVDNCATCRNRHMDLCIECOANQSASSETSECTVAMGVNCHAFH	80
Dd	2 KVKIKCWNGVATWLVAENDENCIGICRMFGCCPDCK-----VPGDCCPLVMGQCSCHFH	56
Qy	81 FHCISRMLKTRQV---CPL 96	
Dd	57 MHCIILKWLHAOCVOOHCPEM 75	

```

RESULT 12
US-09-621-976-5805
; Sequence 5805, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5805

```

```
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -109..-1
US-09-621-976-5805
```

Query Match	30.0%	Score 185;	DB 4;	Length 112;
Best Local Similarity	38.0%;	Pred. No. 7.4e-13;		
Matches	30;	Conservative 13;	Mismatches 28;	Indels 8; Gaps 2;
Qy	21	RFEVKNAVALWANDIVDNCALCRNHIMDLICIQANQASATSECTVAGVGVCHAFH	80	
		: : :           :       : : :		
Db	2	KVKIKNGVATLWLVANDENCIGICRMFNGCCPDCK-----VPGDDCLPLVWGQCSCHCFF	56	
Qy	81	FHCISRWLKTRQV-----CPL	96	
Db	57	MHCILKWLHAAOOVOOHCPM	75	

```

RESULT 13
US-09-621-976-5714
; Sequence 5714, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 5714
; LENGTH: 94
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -53...-1
; NAME/KEY: UNSURE
; LOCATION: 14
; OTHER INFORMATION: xaa = Glu,Gln
US-09-621-976-5714

```

Query Match 17.1%; Score 105.5; DB 4; Length 94;  
Best Local Similarity 36.6%; Pred. No. 0.00031;  
Matches 15; Conservative 8; Mismatches 13; Indels 5; Gaps 1;

RESULT 14  
US-09-949-016-10851  
; Sequence 10851, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10851
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10851

Query Match      14.6%; Score 90; DB 4; Length 195;
Best Local Similarity 26.3%; Pred. No. 0.034;
Matches 25; Conservative 11; Mismatches 15; Indels 44; Gaps 4;

QY 10 PSRTNSGAGK-----KREPVKKNAVALWADIVDNCNCAICRNHIMDLICIEQANQASAT 64
Db 23 PGGSSNGAGESGGQSTFE-----CNICLDTAKDAVI----- 54

QY 65 SEECTVAMGVCNHFHFCISRWLKT---RQVCPL 96
Db 55 -----SLCGHLFCWPCCLHQWLETRPNRQVCPV 81

RESULT 15
US-09-949-016-6249
; Sequence 6249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6249
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6249

Query Match      14.0%; Score 86; DB 4; Length 326;
Best Local Similarity 30.2%; Pred. No. 0.17;
Matches 19; Conservative 11; Mismatches 23; Indels 10; Gaps 2;

QY 43 AICRNHIMDLICIEQANQASATSEECTVAMGVCNHFHFCISRWLKTRQVCPLDNREWE 102
Db 266 AVSRNPLCTLCLE-ERRHPTATP-----CGHLFCWECITAWCSSKAECPLCKRKFP 315

QY 103 FOK 105
Db 316 POK 318

Search completed: June 23, 2005, 15:31:18
Job time : 44 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 15:15:21 ; Search time 162 Seconds  
(without alignments)  
257.840 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVTPSGTNSGAGKK.....KTRQVCPLDNREWFQYKXH 108

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseqp1980a:.\*  
2: Geneseqp1990a:.\*  
3: Geneseqp2000a:.\*  
4: Geneseqp2001a:.\*  
5: Geneseqp2002a:.\*  
6: Geneseqp2003a:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004a:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	108	3	AAB19160 Amino aci
2	616	100.0	108	3	AAB08813 A human c
3	616	100.0	108	3	AAG03890 Human sec
4	616	100.0	108	4	AAB47599 ROC1. 1/2
5	616	100.0	108	5	AAB24620 Human RIN
6	616	100.0	108	7	ADC02481 RING fing
7	616	100.0	108	7	ADP17866 RING fing
8	616	100.0	108	8	ABM82365 Tumour-as
9	616	100.0	108	8	ADSB8291 Human pro
10	534	86.7	108	4	ABB68080 Drosophil
11	514.5	83.5	118	3	AGC23004 Arabidops
12	511	83.0	109	3	AGC23005 Arabidops
13	400.5	65.0	122	4	ABE66109 Drosophi
14	385	62.5	121	3	AAB08814 A yeast c
15	336	54.5	59	8	ADP22560 Sea-squir
16	323	52.4	57	3	AAB41007 Human ORF
17	292	47.4	113	2	AAY06492 Human sen
18	292	47.4	113	3	AAB43295 Human ORF
19	292	47.4	113	8	ADR97346 Human RNF
20	292	47.4	118	4	AAU15873 Human nov
21	292	47.4	118	6	ABU54942 Human nov
22	292	47.4	131	4	AAU16332 Human nov
23	292	47.4	131	6	ABU55401 Human nov
24	291	47.2	113	2	AAY06509 Human sen
25	290.5	47.2	113	2	AAY06491 Mouse sen

26	290.5	47.2	113	3	AAB19161	AAB19161 Amino aci
27	287	46.6	113	4	AAB47600	AAB47600 ROC2. 1/2
28	287	46.6	113	5	AAE24621	AAE24621 Human RIN
29	287	46.6	113	7	ADC02483	ADC02483 RING fing
30	287	46.6	113	7	ADP17868	ADP17868 RING fing
31	285	46.3	113	2	AAY06508	AAY06508 Human sen
32	283	45.9	113	2	AAY06500	AAY06500 Human sen
33	283	45.9	113	2	AAU08501	AAU08501 Human sen
34	282	45.8	113	2	AAY06496	AAY06496 Human sen
35	282	45.8	113	2	AAU06503	AAU06503 Human sen
36	282	45.8	113	2	AAY06502	AAY06502 Human sen
37	282	45.8	113	2	AAY06506	AAY06506 Human sen
38	282	45.8	113	2	AAY06495	AAY06495 Human sen
39	282	45.8	113	2	AAY06499	AAY06499 Human sen
40	282	45.8	113	2	AAY06498	AAY06498 Human sen
41	282	45.8	113	2	AAY06505	AAY06505 Human sen
42	274	44.5	113	2	AAU06504	AAU06504 Human sen
43	272	44.2	113	2	AAY06497	AAY06497 Human sen
44	272	44.2	113	2	AAY06507	AAY06507 Human sen
45	255.5	41.5	100	8	ADP22582	ADP22582 Sea-squir

## ALIGNMENTS

RESULT 1  
AAB19160

ID AAB19160 standard; protein; 108 AA.

XX AAB19160;

XX 19-FEB-2001 (first entry)

DE Amino acid sequence of human ring finger protein ROC1.

KW ROC1; ROC2; cullin; ring finger protein; APC11; APC complex; SCF pathway;  
KW cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;  
KW tumour.

OS Homo sapiens.

PN WO200058472-A2.

XX 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US008592.

PR 31-MAR-1999; 99US-0127261P.

PR 22-NOV-1999; 99US-0166927P.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Xiong Y, Ohta T;

XX WPI; 2000-647235/62.

DR N-PSDB; AAA96882.

XX Novel nucleic acid encoding cullin regulating ring finger proteins,  
PT termed as ROC proteins similar to anaphase-promoting complex 11, for  
PT therapeutic and diagnostic use.

XX Claim 9; Fig 2A; 83pp; English.

XX The present sequence represents a human ROC1 ring finger protein. The  
CC specification also describes human ROC2, ROC1 and ROC2 are similar to  
CC APC11, a subunit of the APC complex. The proteins stimulate cullin  
CC dependent ubiquitin ligase activity. ROC1 functions in vivo as an  
CC essential regulator of CDK inhibitor Sic1 degradation by the SCF  
CC (undefined) pathway. ROC proteins are useful for screening bioactive  
CC agents that interfere with the binding of ROC proteins with cullin  
CC proteins. Pharmaceutical formulations comprising ROC proteins are useful  
CC for diagnostic and therapeutic purposes, preferably for diagnosing and  
CC treating tumours

XX	Sequence 108 AA;	Best Local Similarity 100.0%; Score 616; DB 3; Length 108;	Best Local Similarity 100.0%; Pred. No. 3.8e-64;	Mismatches 0; Indels 0; Gaps 0;
SQ	Query Match	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
	Best Local Similarity 100.0%; Pred. No. 3.8e-64;	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
	Mismatches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
QY	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
DB	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
QY	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108
DB	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108
RESULT 2				
AAB08813				
ID	AAB08813 standard; protein; 108 AA.			
XX				
AC	AAB08813;			
XX				
DT	02-JAN-2001 (first entry)			
DE	A human cullin-interacting RING-H2 finger protein (Rbx1).			
XX				
KW	Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;			
KW	tumour suppressor; carcinoma; Ring box associated carcinoma;			
KW	von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;			
KW	cerebellar hemangioblastoma; hemangioma; retinal angiomas;			
KW	pheochromocytomas.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200050445-A1.			
XX				
PD	31-AUG-2000.			
XX				
PF	25-FEB-2000; 2000WO-US004838.			
XX				
PR	26-FEB-1999; 99US-0121787P.			
XX				
PA	(OKLA-) OKLAHOMA MEDICAL RES FOUND.			
XX				
PI	Conaway JA, Conaway RC, Kamura T;			
XX				
DR	WPI; 2000-572067/53.			
XX	N-PSDB; AAA74978.			
XX				
PT	Cullin interacting RING-H2 finger protein, a component of von Hippel-Lindau tumor suppressor complex and Skp1-Cdc53p-F-box protein (SCF)			
PT	ubiquitin ligase, useful for diagnosing and treating Ring box protein associated carcinomas.			
XX				
PS	Claim 1; Page 34; 37pp; English.			
XX				
CC	The present sequence represents a human cullin-interacting RING-H2 finger protein (Ring box protein), designated Rbx1. The polypeptide is a tumour suppressor. Rbx1 is useful for diagnosing a predisposition of a patient to certain carcinomas. It is also useful for treating Ring box protein associated carcinomas or augmenting metabolically deficient system in animals. Rbx1 is also useful for evaluating the effectiveness of a therapeutic treatment for Ring box associated carcinomas. Rbx1 can be used to screen for agents which augment or inhibit the activity of other cullin-containing ubiquitin ligase and of the VHL (von Hippel-Lindau) complex controlling the conjugation of ubiquitin or ubiquitin-like proteins to various sets of target proteins. Carcinomas which may be treated include renal carcinomas, cerebellar hemangioblastomas and hemangiomas, retinal angiomas and pheochromocytomas			
XX				
SQ	Sequence 108 AA;			
	Query Match	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
	Best Local Similarity 100.0%; Score 616; DB 3; Length 108;	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
	Mismatches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
QY	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
DB	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60	1 MAAAMDVDTPSGTNSGAGKRRFEVKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
QY	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108	61 ASATSECTVAMGVCNHFHFCISRWLKTRQVCPDLNREWEFQKYGH 108

Db 61 ASATSECTVAVGVCNHAHFHCISRMLKTRQVCPDLNREWEFQKYGH 108  
|||||

RESULT 4  
AAB47599  
ID AAB47599 standard; protein; 108 AA.

XX AAB47599;  
DT 07-JAN-2002 (first entry)  
XX ROC1.

XX Assay; ubiquitin ligase; tag1-ubiquitin; E1; E2;  
KW ubiquitin activating enzyme; ubiquitin conjugating enzyme; E3;  
KW ubiquitin ligase; ubiquitination modulator.

XX Unidentified.

XX WO200175145-A2.

XX 11-OCT-2001.

XX 03-APR-2001; 2001WO-US010906.

XX 03-APR-2000; 2000US-00542497.

XX (RIGE-) RIGEL PHARM INC.

XX Issakani SD, Huang J, Sheung J, Pray TR;

XX WPI; 2001-626445/72.

XX Assaying ubiquitin ligase activity for identifying modulators of  
PT ubiquitination, by combining ubiquitin, ubiquitin activating conjugating  
PT enzyme, ubiquitin ligase and measuring amount of ubiquitin bound to the  
PT ligase.

PS Example 1; Fig 11; 98pp; English.

XX The sequences given in AAB47596-602 are proteins which may be used in the  
CC method of the invention for assaying ubiquitin ligase activity. The  
CC method comprises combining under conditions that favour ubiquitin ligase  
CC activity, tag1-ubiquitin, E1 (ubiquitin activating enzyme), E2 (ubiquitin  
CC conjugating enzyme) and E3 (ubiquitin ligase) and measuring the amount of  
CC tag1-ubiquitin bound to the E3. The method is useful for assaying  
CC ubiquitin ligase activity and ubiquitination enzyme activity which is  
CC useful for identifying ubiquitination modulator. The method comprises  
CC combining tag1-ubiquitin, the modulator, E1, E2 and tag2-E3 and measuring  
CC the amount of tag1-ubiquitin bound to tag2-E3 or combining tag1-  
CC ubiquitin, modulator, E1 and tag3-E2 and measuring the amount of tag1-  
CC ubiquitin bound to tag3-E2. Ubiquitin ligase activity is measured  
CC directly, eliminating the need for target proteins and subsequent  
CC analysis such as separating ligated from unligated material in an SDS-  
CC PAGE procedure. This allows multi-well array analysis and high throughput  
CC screening techniques for modulators of ubiquitination activity. The  
CC method also allows the analysis of many different combinations of E3  
CC components and E2/E3 combinations without requiring prior identification  
CC of specific target substrates. Ubiquitin is labeled, directly or  
CC indirectly and this allows for easy and rapid detection and measurement  
CC of ligated ubiquitin

XX Sequence 108 AA;

Query Match 100.0%; Score 616; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.8e-64;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAMVDVTPSGTNSGAGKRFVKKNAVALWAWDVIVDNCAICRNHMDLCTECQANQ 60  
Db 1 MAAMVDVTPSGTNSGAGKRFVKKNAVALWAWDVIVDNCAICRNHMDLCTECQANQ 60  
|||||

Qy 61 ASATSECTVAVGVCNHAHFHCISRMLKTRQVCPDLNREWEFQKYGH 108  
Db 61 ASATSECTVAVGVCNHAHFHCISRMLKTRQVCPDLNREWEFQKYGH 108  
|||||

RESULT 5  
RAE24620  
ID AAE24620 standard; protein; 108 AA.

XX AAE24620;

XX 22-OCT-2002 (first entry)

XX Human RING finger protein, ROC1.

XX Ubiquitin ligase; UL; tag1-ubiquitin; E1; ubiquitin-activating enzyme;  
KW E2; ubiquitin-conjugating enzyme/ubiquitin carrier protein; E3; human;  
KW RING finger protein.

XX Homo sapiens.

XX US2002042083-A1.

XX 11-APR-2002.

XX 03-APR-2001; 2001US-00826312.

XX 03-APR-2000; 2000US-00542497.

XX (RIGE-) RIGEL PHARM INC.

XX Issakani SD, Huang J, Sheung J, Pray TR;

XX WPI; 2002-488718/52.

XX Assay for ubiquitin ligase activity, useful for identifying modulators,  
PT by measuring binding of labeled ubiquitin to ubiquitin ligase.

PS Disclosure; Fig 11; 56pp; English.

XX The invention relates to an assay for ubiquitin ligase (UL) activity  
CC which comprises incubating tag1-ubiquitin, E1 (ubiquitin-activating  
CC enzyme), E2 (ubiquitin-conjugating enzyme/ubiquitin carrier protein) and  
CC E3 (UL) and measuring the amount of tag1-ubiquitin bound to E3. The  
CC method is particularly used to screen for modulators of UL activity. The  
CC present sequence is human RING finger protein, ROC1

XX Sequence 108 AA;

Query Match 100.0%; Score 616; DB 5; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.8e-64;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAMVDVTPSGTNSGAGKRFVKKNAVALWAWDVIVDNCAICRNHMDLCTECQANQ 60  
Db 1 MAAMVDVTPSGTNSGAGKRFVKKNAVALWAWDVIVDNCAICRNHMDLCTECQANQ 60  
|||||

Qy 61 ASATSECTVAVGVCNHAHFHCISRMLKTRQVCPDLNREWEFQKYGH 108  
Db 61 ASATSECTVAVGVCNHAHFHCISRMLKTRQVCPDLNREWEFQKYGH 108  
|||||

RESULT 6  
ADC02481  
ID ADC02481 standard; protein; 108 AA.

XX ADC02481;

XX 18-DEC-2003 (first entry)

XX RING finger protein ROC1.

XX ubiquitin; Mdm2; p53; RING finger protein; ROC1.

```

XX OS Unidentified.
XX PN US2003104474-A1.
XX PD 05-JUN-2003.
XX PF 26-MAR-2002; 2002US-00108767.
XX PR 03-APR-2000; 2000US-00542487.
XX PR 03-APR-2001; 2001US-00826312.
XX PR 04-MAR-2002; 2002US-00091139.
XX PA (ISSA/) ISSAKANI S D.
XX PA (HUAN/) HUANG J.
XX PA (SHEU/) SHEUNG J.
XX PA (PRAY/) PRAY T R.
XX PI Issakani SD, Huang J, Sheung J, Pray TR;
XX WPI; 2003-787056/74.
XX DR Assaying for candidate agent that modulates attachment of ubiquitin
XX PT moiety to Mdm2 protein comprises combining ubiquitin agent comprising
XX PT ubiquitin moiety, Mdm2 protein, and candidate agent.
XX PS Disclosure; Fig 11; 96pp; English.
XX SS The present sequence represents a method of assaying for a candidate
XX CC agent that modulates the attachment of a ubiquitin moiety to an Mdm2
XX CC protein. The method is useful for assaying for a candidate agent that
XX CC modulates the attachment of a ubiquitin moiety to an Mdm2 or p53 protein.
XX CC The present sequence represents the amino acid sequence of RING finger
XX CC protein ROC1.
XX SQ Sequence 108 AA;
    Query Match 100.0%; Score 616; DB 7; Length 108;
    Best Local Similarity 100.0%; Pred. No. 3.8e-64;
    Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMDVDTFSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLCEQANQ 60
DB 1 MAAAMDVDTFSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLCEQANQ 60
QY 61 ASATSEECTVAMGVCNHAFFHCISRWLKTRQVCPLDNRWEFQKYGH 108
DB 61 ASATSEECTVAMGVCNHAFFHCISRWLKTRQVCPLDNRWEFQKYGH 108
RESULT 7
ADF17866
ID ADF17866 standard; protein; 108 AA.
XX AC ADF17866;
XX DT 12-FEB-2004 (first entry)
XX DE RING finger protein ROC1, an E3 ubiquitin ligating agent.
XX KW ROC1; ubiquitin-mediated proteolysis; high-throughput screening;
XX KW E3 ubiquitin ligating agent; RING finger.
XX OS Unidentified.
XX PN US2003108947-A1.
XX PD 12-JUN-2003.
XX PF 20-MAY-2002; 2002US-00152156.
XX PR 03-APR-2000; 2000US-00542497.
XX PR 03-APR-2001; 2001US-00826312.
PR 18-MAY-2001; 2001US-0291836P.
PR 04-MAR-2002; 2002US-00091139.
PR 04-MAR-2002; 2002US-00091174.
PR 26-MAR-2002; 2002US-00108767.
PR 26-MAR-2002; 2002US-00109460.
XX (ISSA/) ISSAKANI S D.
XX PA (HUAN/) HUANG J.
XX PA (SHEU/) SHEUNG J.
XX PA (PRAY/) PRAY T R.
XX PI Issakani SD, Huang J, Sheung J, Pray TR;
XX WPI; 2003-810851/76.
XX DR Assaying agent that modulates attachment of ubiquitin moiety to
XX PT ubiquitin agent, by combining first ubiquitin agent, candidate agent and
XX PT ubiquitin moiety, and assaying agent.
XX PS Disclosure; Fig 11; 95pp; English.
XX SS This invention relates to a novel method for assaying ubiquitin agents
XX CC that are enzymatic components of ubiquitin-mediated proteolysis.
XX CC Specifically, it refers to a method for identifying ubiquitin-activating,
XX CC -conjugating and -ligating agents that work to modulate the attachment of
XX CC a ubiquitin moiety to at least one other ubiquitin agent. The method
XX CC comprises combining a first ubiquitin agent, candidate agent and
XX CC ubiquitin moiety, and subsequently assaying for the attachment of a
XX CC ubiquitin moiety to the first agent. The present invention does not
XX CC require a ubiquitin target protein and allows the analysis of many
XX CC different combinations of ubiquitin agents, without requiring the
XX CC identification of specific target proteins. Consequently, this method is
XX CC easy, applicable for high-throughput screening applications and so can
XX CC rapidly detect the activity of ubiquitin agents. This polypeptide
XX CC sequence is the RING finger protein ROC1, an E3 ubiquitin ligating agent
XX CC of the invention.
XX SQ Sequence 108 AA;
    Query Match 100.0%; Score 616; DB 7; Length 108;
    Best Local Similarity 100.0%; Pred. No. 3.8e-64;
    Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMDVDTFSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLCEQANQ 60
DB 1 MAAAMDVDTFSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLCEQANQ 60
QY 61 ASATSEECTVAMGVCNHAFFHCISRWLKTRQVCPLDNRWEFQKYGH 108
DB 61 ASATSEECTVAMGVCNHAFFHCISRWLKTRQVCPLDNRWEFQKYGH 108
RESULT 8
ABM82365
ID ABM82365 standard; protein; 108 AA.
XX AC ABM82365;
XX DT 18-NOV-2004 (first entry)
XX DE Tumour-associated antigenic target (TAT) polypeptide PRO83289, SEQ:6073.
XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX KW tumour; diagnosis; cell proliferative disorder; breast cancer;
XX KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX KW central nervous system cancer; bladder cancer; pancreatic cancer;
XX KW cervical cancer; melanoma; leukaemia; hybridisation probe;
XX KW chromosome identification; chromosome mapping; gene mapping;
XX KW gene therapy; cytostatic.
XX OS Homo sapiens.
XX PN WO2004030615-A2.

```

XX PD 15-APR-2004.  
XX PF 29-SEP-2003; 2003WO-US028547.  
XX PR 02-OCT-2002; 2002US-0414971P.  
XX PA (GETH ) GENENTECH INC.  
XX PI Wu TD, Zhang Z, Zhou Y;  
XX DR WPI; 2004-347921/32.  
XX DR N-PSDB; ACN40951.  
XX PT New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX PS Claim 12; SEQ ID NO 6073; 7273pp; English.  
XX CC The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT polypeptide of the invention  
XX SQ Sequence 108 AA;  
Query Match 100.0%; Score 616; DB 8; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.8e-64;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAAMDVDTPSGTNSGAGKRFVKKNNAVALWAWDIIVVNDCAICRNHIMDLCTECQANQ 60  
DB 1 MAAAMDVDTPSGTNSGAGKRFVKKNNAVALWAWDIIVVNDCAICRNHIMDLCTECQANQ 60  
QY 61 ASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFKYGH 108  
DB 61 ASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFKYGH 108  
RESULT 9  
ADS88291  
ID ADS88291 standard; protein; 108 AA.  
XX AC ADS88291;  
XX DT 18-NOV-2004 (first entry)  
XX DE Human protein of a TNF-alpha signalling pathway protein complex Seq 146.  
XX KW protein complex; tumour necrosis factor-alpha signalling pathway;  
KW TNF-alpha; chronic inflammatory disease; rheumatoid arthritis;  
KW inflammatory bowel disease; infectious disease; septic shock;  
KW bacterial infection; neurological disease; stroke-induced inflammation;  
KW neurodegenerative disease; cancer; antiinflammatory; antiarthritic;  
KW antirheumatic; cytostatic; antibacterial; gene therapy; human.

XX OS Homo sapiens.  
XX PN WO2004035783-A2.  
XX PD 29-APR-2004.  
XX PF 24-SEP-2003; 2003WO-EP050655.  
XX PR 26-SEP-2002; 2002EP-00021809.  
XX PR 10-FEB-2003; 2003EP-00100274.  
XX PA (CELL-) CELLZONE AG.  
XX PI Bouwmeester T, Huhse B, Bauch A, Ruffner H, Bauer A, Kuester B;  
PI Superti-Furga G, Kruse U;  
XX WPI; 2004-348460/32.  
XX PT New protein complex comprising at least one first and second protein of  
PT the Tumor Necrosis Factor-alpha (TNF-alpha)-signaling pathway, useful for  
PT diagnosing or treating inflammation, neurological diseases, infectious  
PT diseases or cancer.  
XX PS Example; SEQ ID NO 146; 1980pp; English.  
XX CC This invention relates to novel protein complexes of the tumour necrosis  
CC factor-alpha (TNF-alpha) signalling pathway. Specifically, it refers to  
CC methods for preparing these complexes comprising at least two component  
CC proteins, as well as screening methods to identify modulators of the  
CC pathway, which include antibodies, agonists and antagonists thereof. The  
CC present invention describes a protein complex and kit that are useful for  
CC diagnosing, prognosing or treating chronic inflammatory diseases such as  
CC rheumatoid arthritis and inflammatory bowel disease; infectious diseases  
CC such as septic shock and bacterial infections; neurological diseases such  
CC as stroke-induced inflammation in neurons; neurodegenerative diseases and  
CC cancer. Accordingly, these complexes can be used for the development of  
CC pharmaceutical compositions that exhibit antiinflammatory, antiarthritic,  
CC antirheumatic, cytostatic and antibacterial activities and can be used  
CC for gene therapy purposes. In particular, the invention further provides  
CC siRNA-oligonucleotides useful for inhibiting protein expression for in  
CC vitro or cell culture assays. This polypeptide is a human protein that  
CC can be used in combination with other proteins provided in the  
CC specification to form novel complexes of the TNF-alpha signalling pathway  
CC of the invention.  
XX SQ Sequence 108 AA;  
Query Match 100.0%; Score 616; DB 8; Length 108;  
Best Local Similarity 100.0%; Pred. No. 3.8e-64;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAAMDVDTPSGTNSGAGKRFVKKNNAVALWAWDIIVVNDCAICRNHIMDLCTECQANQ 60  
DB 1 MAAAMDVDTPSGTNSGAGKRFVKKNNAVALWAWDIIVVNDCAICRNHIMDLCTECQANQ 60  
QY 61 ASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFKYGH 108  
DB 61 ASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFKYGH 108  
RESULT 10  
ABB68080  
ID ABB68080 standard; protein; 108 AA.  
XX AC ABB68080;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 31032.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

```
XX Drosophila melanogaster.
OS WO200171042-A2.
PN
XX
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
XX N-PSDB; ABL12183.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PT
XX
XX Disclosure; SEQ ID NO 31032; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 108 AA;
SQ
Query Match 86.7%; Score 534; DB 4; Length 108;
Best Local Similarity 87.2%; Pred. No. 1.6e-54;
Matches 95; Conservative 4; Mismatches 4; Indels 6; Gaps 2;
QY 5 MDYD-----TPSGTNSGAGKKRFEVKKNAVALWAWDIIVDNCAICRNHIMDLICICQAN 59
DB 1 MEVDEDCGYEVPSSSSKG-DKREFEVKKNAVALWAWDIIVDNCAICRNHIMDLICICQAN 59
QY 60 QASATSEECTVANGVNCNHAFPHFCISRWLKTRQVCPDLNREWFQKYGH 108
DB 60 QASATSEECTVANGVNCNHAFPHFCISRWLKTRQVCPDLNREWFQKYGH 108
RESULT 11
AAG23004
ID AAG23004 standard; protein; 118 AA.
XX
XX AAG23004;
AC
XX
XX 17-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 26148.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX
```

PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158023P.
PR 17-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 18-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160814P.
PR 23-JUL-1999;	99US-0145145P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160981P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0160989P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0147302P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0147303P.	PR 26-OCT-1999;	99US-0161359P.
PR 28-JUL-1999;	99US-0147304P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0147305P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0147306P.	PR 26-OCT-1999;	99US-0161362P.
PR 02-AUG-1999;	99US-0147307P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147308P.	PR 28-OCT-1999;	99US-0161922P.
PR 04-AUG-1999;	99US-0147309P.	PR 28-OCT-1999;	99US-0161933P.
PR 04-AUG-1999;	99US-0147310P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147311P.		
PR 05-AUG-1999;	99US-0147312P.		
PR 06-AUG-1999;	99US-0147313P.		
PR 06-AUG-1999;	99US-0147314P.		
PR 09-AUG-1999;	99US-0147315P.		
PR 09-AUG-1999;	99US-0147316P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148311P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149428P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 25-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 30-AUG-1999;	99US-0151080P.		
PR 31-AUG-1999;	99US-0151303P.		
PR 01-SEP-1999;	99US-0151438P.		
PR 07-SEP-1999;	99US-0151930P.		
PR 10-SEP-1999;	99US-0152363P.		
PR 13-SEP-1999;	99US-0153070P.		
PR 15-SEP-1999;	99US-0153758P.		
PR 16-SEP-1999;	99US-0154018P.		
PR 20-SEP-1999;	99US-0154039P.		
PR 22-SEP-1999;	99US-0154779P.		
PR 23-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156598P.		

Query Match 83.5%; Score 514.5; DB 3; Length 118;

Best Local Similarity 78.6%; Pred. No. 3.5e-52; Indels 11; Gaps 2; Matches 92; Conservative 6; Mismatches 8;

Qy	3	AAMDVDT---PSG-----TNSGAGKGRPEVKKNNAVALWWDIVVDNCAICRNHMD 51
Db	2	ATLDSVTMTIPAGASSVAASSNNKKAKRFEIKKSAVALWWDIVVDNCAICRNHMD 61
Qy	52	LCIECQANQASATSECTVAVGVCNHAHAFHFCISRWLKTRQVCLDNRWFQKYGH 108
Db	62	LCIECQANQASATSECTVAVGVCNHAHAFHFCISRWLKTRQVCLDNRWFQKYGH 118

RESULT 12

AAG23005 ID AAG23005 standard; protein; 109 AA.

XX AAG23005; AC AAG23005; DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 26149.

XX Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131144P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134376P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139751P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139859P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-01421154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142330P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 20-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149910P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.





PS Disclosure; Page 34-35; 37pp; English.

XX The present sequence represents a yeast cullin-interacting RING-H2 finger

CC protein (RING box protein), designated Rbx1. The human Rbx1 polypeptide

CC is a tumour suppressor. Human Rbx1 is useful for diagnosing a

CC predisposition of a patient to certain carcinomas. It is also useful for

CC treating RING box protein associated carcinomas or augmenting

CC metabolically deficient system in animals. Human Rbx1 is also useful for

CC evaluating the effectiveness of a therapeutic treatment for RING box

CC associated carcinomas. Human Rbx1 can be used to screen for agents which

CC augment or inhibit the activity of other cullin-containing ubiquitin

CC ligase and of the VHL (von Hippel-Lindau) complex controlling the

CC conjugation of ubiquitin or ubiquitin-like proteins to various sets of

CC target proteins. Carcinomas which may be treated include renal

CC carcinomas, cerebellar hemangioblastomas and hemangiomas, retinal

CC angiomas and pheochromocytomas

XX

SQ Sequence 121 AA;

Query Match 62.5%; Score 385; DB 3; Length 121;

Best Local Similarity 56.6%; Pred. No. 5.7e-37;

Matches 64; Conservative 10; Mismatches 29; Indels 10; Gaps 1;

QY 5 MDVDPSTGNSGAG-----KKRFVKKNAVALMWDIVDNCACIRNHIMDLCI 54

DB 8 MDVDEDSQNIQSSNQAPVETKKRFEIKKWTAVAFWSWDIAVDNCAICRNHIMEPCI 67

QY 55 EQQANQASATSECTVANGVGNHAFHFCISRWLKTRQVCPDLNREWEFQKYG 107

DB 68 EQPKAMTDTDNECVAAGVGNHAFHFCINRKWIKTRDACPLDNPQWLARCG 120

RESULT 15

ADP22560

ID ADP22560 standard; protein; 59 AA.

XX

AC ADP22560;

XX

DT 12-AUG-2004 (first entry)

XX

DE Sea-squirt (Ciona intestinalis) zinc finger protein #67.

XX

KW sea-squirt; zinc finger protein; gene detection; drug development;

KW zinc finger protein-associated disease.

XX

OS Ciona intestinalis.

XX

PN JF2004057126-A.

XX

PD 26-FEB-2004.

XX

PF 31-JUL-2002; 2002JP-00222484.

XX

PR 31-JUL-2002; 2002JP-00222484.

XX

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX

DR WPI; 2004-208711/20.

DR N-PSDB; ADP22559.

XX Novel gene encoding zinc finger protein, useful as probe in gene

PT detecting instruments and in development of drug for treating zinc finger

PT protein associated diseases.

XX

PS Claim 1; SEQ ID NO 134; 972pp; Japanese.

XX

CC The invention comprises the amino acid and coding sequences of sea-squirt

CC (Ciona intestinalis) zinc finger proteins. The DNA and protein sequences

CC of the invention are useful in a gene detecting instrument. The DNA and

CC protein sequences of the invention are useful in the development of drugs

CC for the treatment of zinc finger protein-associated diseases. The present

CC amino acid sequence represents a sea-squirt zinc finger protein of the

CC invention.

XX

SQ Sequence 59 AA;

Query Match 54.5%; Score 336; DB 8; Length 59;

Best Local Similarity 98.3%; Pred. No. 1.4e-31;

Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 50 MDLCIECQANQASATSECTVANGVGNHAFHFCISRWLKTRQVCPDLNREWEFQKYG 107

DB 1 MDLCIECQANQASATSECTVANGVGNHAFHFCISRWLKTRQVCPDLNREWEFQKYG 58

Search completed: June 23, 2005, 15:26:42

Job time : 165 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: June 23, 2005, 15:29:52 ; Search time 158 Seconds  
(without alignments)  
262.856 Million cell updates/sec

Title: US-09-541-462B-2

Perfect score: 616  
Sequence: 1 MAAMDVDTSGTSGAGKK.....KTRQVCPLDREWEFKYGH 108

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	108	9	US-09-826-312-6
2	616	100.0	108	14	US-10-108-767-6
3	616	100.0	108	14	US-10-152-156-6
4	616	100.0	108	17	US-10-913-937-1
5	616	100.0	108	17	US-10-835-096-6
6	534	86.7	108	17	US-10-913-937-6
7	524	85.1	124	15	US-10-767-701-43736
8	518.5	84.2	118	15	US-10-424-599-148916
9	512	83.1	109	16	US-10-425-115-201494
10	512	83.1	111	16	US-10-767-701-43735
11	506.5	82.2	152	15	US-10-424-599-221431

12	493	80.0	110	17	US-10-913-937-7	Sequence 7, Appli
13	420.5	68.3	106	15	US-10-424-599-148915	Sequence 148915,
14	420	68.2	82	16	US-10-425-115-191492	Sequence 191492,
15	420	68.2	97	16	US-10-425-115-291051	Sequence 291051,
16	385	62.5	121	17	US-10-913-937-2	Sequence 2, Appli
17	356.5	57.9	98	16	US-10-425-115-352839	Sequence 352839,
18	313	50.8	75	15	US-10-424-599-230014	Sequence 230014,
19	308.5	50.1	132	16	US-10-425-115-336213	Sequence 336213,
20	294.5	47.8	109	16	US-10-767-701-51814	Sequence 51814, A
21	292	47.4	118	9	US-09-764-864-826	Sequence 826, App
22	292	47.4	131	9	US-09-764-864-1285	Sequence 1285, Ap
23	287	46.6	64	15	US-10-424-599-238320	Sequence 238320,
24	287	46.6	113	9	US-09-826-312-8	Sequence 8, Appli
25	287	46.6	113	14	US-10-108-767-8	Sequence 8, Appli
26	287	46.6	113	14	US-10-152-156-8	Sequence 8, Appli
27	287	46.6	113	17	US-10-835-096-8	Sequence 8, Appli
28	282	45.8	68	15	US-10-424-599-242288	Sequence 242288,
29	228	37.0	40	15	US-10-424-599-264079	Sequence 264079,
30	213	34.6	88	9	US-09-925-298-620	Sequence 620, App
31	213	34.6	88	14	US-10-102-806-620	Sequence 620, App
32	213	34.6	88	15	US-10-276-774-2678	Sequence 2678, Ap
33	213	34.6	91	15	US-10-264-049-2937	Sequence 2937, Ap
34	213	34.6	105	9	US-09-764-864-1274	Sequence 1274, Ap
35	210	34.1	124	9	US-09-764-864-1284	Sequence 1284, Ap
36	208	33.8	84	11	US-09-978-360A-409	Sequence 409, App
37	208	33.8	84	15	US-10-315-664-77	Sequence 77, Appl
38	208	33.8	84	15	US-10-221-625-94	Sequence 94, Appl
39	208	33.8	100	15	US-10-424-599-253860	Sequence 253860,
40	203	33.0	84	16	US-10-767-701-51781	Sequence 51781, A
41	202	32.8	84	9	US-09-826-312-5	Sequence 5, Appli
42	202	32.8	84	14	US-10-108-767-5	Sequence 5, Appli
43	202	32.8	84	14	US-10-152-156-5	Sequence 5, Appli
44	202	32.8	84	17	US-10-835-096-5	Sequence 5, Appli
45	196	31.8	85	15	US-10-264-049-2926	Sequence 2926, Ap

## ALIGNMENTS

### RESULT 1

US-09-826-312-6  
; Sequence 6, Application US/09826312  
; Patent No. US20020042083A1  
; GENERAL INFORMATION:  
; APPLICANT: Issakani, Sarkiz D.  
; APPLICANT: Huang, Jianing  
; APPLICANT: Sheung, Julie  
; APPLICANT: Pray, Todd R.  
; TITLE OF INVENTION: UBIQUITIN LIGASE ASSAY  
; FILE REFERENCE: A-68613-1/RMS/JJD  
; CURRENT APPLICATION NUMBER: US/09/826,312  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 09/542,497  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-826-312-6

Query Match	100.0%	Score 616;	DB 9;	Length 108;
Best Local Similarity	100.0%;	Pred. No. 7.7e-60;		
Matches 108;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAAMDVDTSGTSGAGKKFEVKKNNAVALWAWDIWDNCAICRNHIMDLCEQANQ	60	
Db	1	MAAMDVDTSGTSGAGKKFEVKKNNAVALWAWDIWDNCAICRNHIMDLCEQANQ	60	
Qy	61	ASATSEECTVAGVNCNHAFFHCISRLWKTQVCPLDREWEFKYGH	108	
Db	61	ASATSEECTVAGVNCNHAFFHCISRLWKTQVCPLDREWEFKYGH	108	

```
RESULT 2
US-10-108-767-6
; Sequence 6, Application US/10108767
; Publication No. US2003010474A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: ASSAYS FOR IDENTIFYING UBIQUITIN AGENTS AND FOR IDENTIFYING AGENTS
; FILE OF INVENTION: MODIFY THE ACTIVITY OF UBIQUITIN AGENTS
; FILE REFERENCE: A-68613-5/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/108,767
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 10/091,139
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-767-6

Query Match      100.0%; Score 616; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
Db 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60

QY 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108
Db 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 3
US-10-152-156-6
; Sequence 6, Application US/10152156
; Publication No. US20030108947A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: ASSAYS FOR IDENTIFYING UBIQUITIN AGENTS AND FOR IDENTIFYING AGENTS
; FILE OF INVENTION: MODIFY THE ACTIVITY OF UBIQUITIN AGENTS
; FILE REFERENCE: A-68613-6/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/152,156
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 10/091,174
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 10/091,139
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 10/109,460
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 10/108,767
; PRIOR FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/291,836
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
US-10-152-156-6

Query Match      100.0%; Score 616; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
Db 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60

QY 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108
Db 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 4
US-10-913-937-1
; Sequence 1, Application US/10913937
; Publication No. US20050019813A1
; GENERAL INFORMATION:
; APPLICANT: Conaway, Joan A.
; APPLICANT: Kamura, Takumi
; APPLICANT: Okanoma Medical Research Foundation
; TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor
; FILE OF INVENTION: Complex and SCF Ubiquitin Ligase
; FILE REFERENCE: 021044-004600US
; CURRENT APPLICATION NUMBER: US/10/913,937
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US/09/914,324
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/121,787
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: WO PCT/US00/04838
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ring box protein 1 (Rbx1)
US-10-913-937-1

Query Match      100.0%; Score 616; DB 17; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
Db 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60

QY 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108
Db 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 5
US-10-835-096-6
; Sequence 6, Application US/10835096
; Publication No. US20050032139A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
```

```
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-152-156-6

Query Match      100.0%; Score 616; DB 14; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
Db 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60

QY 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108
Db 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 4
US-10-913-937-1
; Sequence 1, Application US/10913937
; Publication No. US20050019813A1
; GENERAL INFORMATION:
; APPLICANT: Conaway, Joan A.
; APPLICANT: Kamura, Takumi
; APPLICANT: Okanoma Medical Research Foundation
; TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor
; FILE OF INVENTION: Complex and SCF Ubiquitin Ligase
; FILE REFERENCE: 021044-004600US
; CURRENT APPLICATION NUMBER: US/10/913,937
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US/09/914,324
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/121,787
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: WO PCT/US00/04838
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human ring box protein 1 (Rbx1)
US-10-913-937-1

Query Match      100.0%; Score 616; DB 17; Length 108;
Best Local Similarity 100.0%; Pred. No. 7,7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60
Db 1 MAAAMDVTPSGTNSGAGKKRFEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQANQ 60

QY 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108
Db 61 ASATSEECTVAMGVCNHAHFHCHISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 5
US-10-835-096-6
; Sequence 6, Application US/10835096
; Publication No. US20050032139A1
; GENERAL INFORMATION:
; APPLICANT: Issakani, Sarkiz D.
; APPLICANT: Huang, Jianing
; APPLICANT: Sheung, Julie
; APPLICANT: Pray, Todd R.
; TITLE OF INVENTION: Ubiquitin Ligase Assay
```

```

; FILE REFERENCE: 021044-007010US
; CURRENT APPLICATION NUMBER: US/10/835,096
; CURRENT FILING DATE: 2004-04-28
; PRIOR APPLICATION NUMBER: US/09/826,312
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 09/542,497
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: RING finger protein ROC1
US-10-835-096-6

Query Match      100.0%; Score 616; DB 17; Length 108;
Best Local Similarity 100.0%; Pred. No. 7.7e-60;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPSGTNSGAGKRFVKKNAVALWAWDIIVDNCACRHNHMDLCIECOANQ 60
DB 1 MAAAMDVTPSGTNSGAGKRFVKKNAVALWAWDIIVDNCACRHNHMDLCIECOANQ 60

QY 61 ASATSEECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108
DB 61 ASATSEECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108

RESULT 6
US-10-913-937-6
; Sequence 6, Application US/10913937
; Publication No. US20050019813A1
; GENERAL INFORMATION:
; APPLICANT: Conaway, Joan A.
; APPLICANT: Conaway, Ronald C.
; APPLICANT: Kamura, Takumi
; TITLE OF INVENTION: Oklahoma Medical Research Foundation
; TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor
; FILE REFERENCE: 021044-004600US
; CURRENT APPLICATION NUMBER: US/10/913,937
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US/09/914,324
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/121,787
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: WO PCT/US00/04838
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Drosophila melanogaster ring box protein 1 (Rbx1)
US-10-913-937-6

Query Match      86.7%; Score 534; DB 17; Length 108;
Best Local Similarity 87.2%; Pred. No. 8.1e-51;
Matches 95; Conservative 4; Mismatches 4; Indels 6; Gaps 2;

QY 5 MDVD-----TPSGTNSGAGKRFVKKNAVALWAWDIIVDNCACRHNHMDLCIECOAN 59
DB 1 MEVDDGGEVSSSSKSG-DKKRFVKKNAVALWAWDIIVDNCACRHNHMDLCIECOAN 59

QY 60 QASATSEECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108
DB 60 QASATSEECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108
```

```

RESULT 7
US-10-767-701-43736
; Sequence 43736, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43736
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB5052-011-A1-XP1-A5.pep
US-10-767-701-43736

Query Match      85.1%; Score 524; DB 16; Length 124;
Best Local Similarity 90.9%; Pred. No. 1.2e-49;
Matches 90; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 10 PSGTNSGAGKRFVKKNAVALWAWDIIVDNCACRHNHMDLCIECOANQASATSECT 69
DB 26 PSSSTRKPKRFEIKKNAVALWAWDIIVDNCACRHNHMDLCIECOANQASATSECT 85

QY 70 VAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108
DB 86 VAVGVCNHAHFHCISRWLKTRQVCPLDNSEWEFQKYGH 124

RESULT 8
US-10-424-599-148916
; Sequence 148916, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 148916
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_105494C.1.pep
US-10-424-599-148916

Query Match      84.2%; Score 518.5; DB 15; Length 118;
Best Local Similarity 80.3%; Pred. No. 4.5e-49;
Matches 94; Conservative 5; Mismatches 7; Indels 11; Gaps 3;

QY 3 AAMDVDT---PSG-TNSGAG-----KKRFEVKNAVALWAWDIIVDNCACRHNHMD 51
DB 2 ATLDSDVTVPAGEASSSAGPSSSTKKPKRFEIKKNAVALWAWDIIVDNCACRHNHMD 61

QY 52 LCICCOANQASATSECTVAVGVCNHAHFHCISRWLKTRQVCPLDNREWEFQKYGH 108
DB 62 LCICCOANQASATSECTVAVGVCNHAHFHCISRWLKTRQVCPLDNSEWEFQKYGH 118

RESULT 9
US-10-425-115-201494
```



RESULT 13  
US-10-424-599-148915  
; Sequence 148915, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 148915  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(106)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT MRT3847\_105493C.1.pep  
US-10-424-599-148915

Query Match 68.3%; Score 420.5; DB 15; Length 106;  
Best Local Similarity 69.2%; Pred. No. 2.4e-38;  
Matches 81; Conservative 5; Mismatches 8; Indels 23; Gaps 4;  
QY 3 AAMDVDT---PSG-TNSGAG-----KCRFEVKKNNAVALWAWDIVDNCALCRNHMD 51  
DB 2 ATLSDVTVFPAGEASSSGSSSTKPKRFEIKKNNAVSLWAWDIVDNCALCRNHMD 61  
QY 52 LCIEQANQASATSECTVANGVGNHAFHFHCISRWLKTRQVCPDLNREWEFKYGH 108  
DB 62 LCIEQANQASATSECTVANGVGNHAFHFHCISRWLKTRQVCPDLNREWEFKYGH 106

RESULT 14  
US-10-425-115-191492  
; Sequence 191492, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 191492  
; LENGTH: 82  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_10621C.1.pep  
US-10-425-115-191492

Query Match 68.2%; Score 420; DB 16; Length 82;  
Best Local Similarity 98.6%; Pred. No. 2.1e-38;  
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 36 DIWDNCAICRNHMDLCIEQANQASATSECTVANGVGNHAFHFHCISRWLKTRQVCP 95  
DB 10 DIVDNCALCRNHMDLCIEQANQASATSECTVANGVGNHAFHFHCISRWLKTRQVCP 69

QY 96 LDNREWEFKYGH 108  
DB 70 LDNREWEFKYGH 82  
RESULT 15  
US-10-425-115-291051  
; Sequence 291051, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 291051  
; LENGTH: 97  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(97)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_28527C.1.pep  
US-10-425-115-291051  
Query Match 68.2%; Score 420; DB 16; Length 97;  
Best Local Similarity 80.9%; Pred. No. 2.5e-38;  
Matches 72; Conservative 7; Mismatches 10; Indels 0; Gaps 0;  
QY 20 KRFEVKKNNAVALWAWDIVDNCALCRNHMDLCIEQANQASATSECTVANGVGNHAF 79  
DB 2 ERCKIKKNVSVSLWAWHVRDSCALCRNHMDLCIEQANQASATSECTVANGVGNHAF 61  
QY 80 HFHCISRWLKTRQVCPDLNREWEFKYGH 108  
DB 62 HFHCISRWLKTRQVCPDLNREWEFKYGH 90  
Search completed: June 23, 2005, 15:43:29  
Job time : 160 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 15:19:51 ; Search time 39 Seconds  
(without alignments)  
266.446 Million cell updates/sec

Title: US-09-541-462B-2

Perfect score: 616

Sequence: 1 MAAAMDVDTSGTNSGAGKK.....KTRQVCPLDNREWEFOKYGH 108

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79.\*

2: Pirl.\*

3: Pirl3.\*

4: Pirl4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	108	2 T51146	ring-box protein 1
2	510	82.8	136	2 T13388	hypothetical prote
3	499	81.0	115	2 T47341	ring-box protein-1
4	493	80.0	110	2 T27823	hypothetical prote
5	469	76.1	107	2 T38310	pop-interacting pr
6	385	62.5	121	2 S66830	hypothetical prote
7	378	61.4	166	2 T18513	hypothetical prote
8	266	43.2	112	2 T29620	hypothetical prote
9	164.5	26.7	94	2 T38652	hypothetical prote
10	146.5	23.8	135	2 T21802	hypothetical prote
11	136	22.1	165	2 S2511	hypothetical prote
12	132.5	21.5	99	2 G30113	Rp1 protein [limp
13	116.5	18.9	2160	2 T20241	hypothetical prote
14	104	16.9	349	2 T06680	hypothetical prote
15	102.5	16.6	249	2 C96775	probable RING zinc
16	102	16.6	210	2 T45654	zinc-finger-like p
17	100	16.2	676	2 T47637	hypothetical prote
18	97	15.7	159	2 T48209	hypothetical prote
19	97	15.7	327	2 D86474	probable RING zinc
20	96	15.6	332	2 E86448	hypothetical prote
21	95.5	15.5	441	2 F71425	hypothetical prote
22	95	15.4	253	2 T06113	hypothetical prote
23	95	15.4	322	2 H5474	hypothetical prote
24	95	15.4	336	2 T28358	ORF MSV197 tryptop
25	95	15.4	362	2 T51464	RING-H2 zinc finger
26	95	15.4	530	2 T50499	hypothetical prote
27	95	15.4	689	2 F84811	probable retroelem
28	92	14.9	185	2 T51844	RING-H2 finger pro
29	91.5	14.9	190	2 T51859	

Query Match

82.8%, Score 510; DB 2; Length 136;

#### RESULT 1

T51146

ring-box protein 1 [imported] - human

C:Species: Homo sapiens (man)

C>Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 09-Jul-2004

C:Accession: T51146

R:Kamura, T.; Keopp, D.M.; Conrad, M.N.; Skowrya, D.; Moreland, R.J.; Iliopoulos, O.; Lar

Science 284, 657-661, 1999

A:Title: Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase.

A:Reference number: 225317; MUID:99234320; PMID:10213691

A:Accession: T51146

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-108 <KAM>

A:Cross-references: UNIPROT:Q8N6Z8; EMBL:AF140598; PIDN:AAD29715.1

C:Genetics:

A:Gene: RBX1

Query Match

Best Local Similarity 100.0%; Score 616; DB 2; Length 108;

Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAAMDVDTSGTNSGAGKKRFEVKWNAVALWAWDIIVVDNCAICRNHIMDLCECQANQ 60

Db 1 MAAAMDVDTSGTNSGAGKKRFEVKWNAVALWAWDIIVVDNCAICRNHIMDLCECQANQ 60

Qy 61 ASATSECTVAMGVCNHAHPCISRWLKTQVCPLDNREWEFOKYGH 108

Db 61 ASATSECTVAMGVCNHAHPCISRWLKTQVCPLDNREWEFOKYGH 108

#### RESULT 2

T13388

hypothetical protein 115C2.11 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004

C:Accession: T13388

R:Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D.

submitted to the EMBL Data Library, May 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: 217665

A:Accession: T13388

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-136 <CAT>

A:Cross-references: UNIPROT:Q9W5E1; EMBL:AL031581; PIDN:CAA20888.1

C:Genetics:

A:Cross-references: FlyBase:FBgn0020381

A:Introns: 64/1

A>Note: EG:115C2.11

Query Match

82.8%, Score 510; DB 2; Length 136;

hypothetical prote  
probable PHD-type  
ORF MSV205 tryptop  
zinc finger protei  
hypothetical prote  
related to COp1-in  
hypothetical prote  
hypothetical prote  
probable RING zinc  
hypothetical prote  
hypothetical prote  
probable RING zinc  
protein F12M16.10  
intermediate early  
protein ZK637.14 [R  
RING-H2 finger pro

#### ALIGNMENTS



A:Cross-references: SGD:S0005493

A:Map position: 15L

A:Note: YOL133w

C:Superfamily: yeast hypothetical protein YOL133w; RING finger homology

Query Match 62.5%; Score 385; DB 2; Length 121;

Best Local Similarity 56.6%; Pred. No. 5.3e-34;

Matches 64; Conservative 10; Mismatches 29; Indels 10; Gaps 1;

Qy 5 MDVTPSGTNSGAG-----KKRFVKKNNAVALWAWDIVVDNCAICRNHIMDLICI 54

Db 8 MDVDEDSQNTAQSSNQAPVETKKRFEIKKWTAVAFWMDIAVDNCAICRNHIMEPCI 67

Qy 55 EQQANQASATSECTVANGVCHHAFHFCISRWLKTRQVCPDLNREWFQKVG 107

Db 68 ECQPKAMTDTNECVAAVGVCHHAFHFCINRWIKTRDACPDLNQWOLARCG 120

#### RESULT 7

T18513 hypothetical protein C0845c - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T18513

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18513

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-166 <LAW>

A:Cross-references: UNIPROT:O77367; EMBL:Z98551; PIDN:CAB11123.2

C:Genetics:

A:Map position: 3

A:Introns: 19/1

A:Note: C0845c

Query Match 61.4%; Score 378; DB 2; Length 166;

Best Local Similarity 59.6%; Pred. No. 4e-33;

Matches 64; Conservative 5; Mismatches 15; Indels 8; Gaps 1;

Qy 22 FEVKKNAVALWAWDIVVDNCAICRNHIMDLICIEQQA-----NQASATSECTVANG 73

Db 72 FKTHKSAVAAWSWDISVDNCAICRNHIMDLICIEQQAQKTTHDNDKXKIDKEGCTVANG 131

Qy 74 VCNHAFHFCISRWLKTRQVCPDLNREWFQK 105

Db 132 VCNHAFHLHLCISRWIKARQVCPDLNRTWEFQK 163

#### RESULT 8

T29620 hypothetical protein R10A10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T29620

R:Wamsley, P.; Bradshaw, H.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans cosmid R10A10.

A:Reference number: Z20653

A:Accession: T29620

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-112 <WAM>

A:Cross-references: UNIPROT:P91404; EMBL:U80449; PIDN:AAB37826.1; GSPDB:GN00019; CESP:R1

A:Experimental source: strain Bristol N2; clone R10A10

C:Genetics:

A:Gene: CESP:R10A10.2

A:Map position: 1

A:Introns: 17/2; 59/1

C:Superfamily: yeast hypothetical protein YOL133w; RING finger homology

Query Match 43.2%; Score 266; DB 2; Length 112;

Best Local Similarity 41.3%; Pred. No. 2.5e-21;

Matches 45; Conservative 20; Mismatches 32; Indels 12; Gaps 2;

Qy 1 MAAAMDYDTPSGTNSGAGKKR-----FEVKKNAVALWAWDIVVDNCAICRNHIMDL 52

Db 1 MNSSNADSQEGSTSAQOKTANPSESPPFVLKKNALAVWAWDECDTCAICRVHLMEE 60

Qy 53 CIEQANQASATSECTVANGVCHHAFHFCISRWLKTRQVCPDLNREW 101

Db 61 CLRCQ-----SPPSAECYVVMGDCNHSFHHCCMTQIRQNNRCPLCQKDW 105

#### RESULT 9

T38652

hypothetical protein SPAC343.03 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: T38652

R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21804

A:Accession: T38652

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-94 <MUR>

A:Cross-references: UNIPROT:Q9UT86; EMBL:AL109739; PIDN:CAB52266.1; GSPDB:

A:Experimental source: strain 972h-; cosmid c343

C:Genetics:

A:Gene: SPDB:SPAC343.03

A:Map position: 1

A:Introns: 7/1; 48/1

Query Match 26.7%; Score 164.5; DB 2; Length 94;

Best Local Similarity 33.7%; Pred. No. 1.5e-10;

Matches 28; Conservative 16; Mismatches 30; Indels 9; Gaps 3;

Qy 21 RFEVKKNAVALWAWDIVVDN-CAICRNHIMDLICIEQANQASATSECTVANGVCHNAP 79

Db 2 KVKILRYHAIANTWTMDTPKDDVCGICRVFPFGCCPQC-----TSPGDCNCPVWGCKRHP 56

Qy 80 HFHCISRWLKT---RQVCPDLNR 99

Db 57 HAHCIQWLATSGSQGQCPMDRQ 79

#### RESULT 10

T21802

hypothetical protein F35G12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T21802

R:Chui, C.

submitted to the EMBL Data Library, October 1994

A:Reference number: Z19473

A:Accession: T21802

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-135 <WIL>

A:Cross-references: UNIPROT:Q20052; EMBL:Z46242; PIDN:CAA86328.1; GSPDB:GN00021; CESP:F3;

A:Experimental source: clone F35G12

C:Genetics:

A:Gene: CESP:F35G12.9

A:Map position: 3

A:Introns: 33/3; 57/3

C:Superfamily: yeast hypothetical protein YOL133w; RING finger homology

F:68-133/Domain: RING finger homology <RRN>

Query Match 23.8%; Score 146.5; DB 2; Length 135;

Best Local Similarity 33.3%; Pred. No. 1.8e-08;

Matches 28; Conservative 10; Mismatches 35; Indels 11; Gaps 3;

Qy 24 VKKNNAVALWAW-DIVVDNCAICRNHIMDLICIEQANQASATSECTVANGVCHNAPFH 82

Db 53 VKKLHVGCEKWLQGGEDTCGICRMFEFESAC-----NMCKPFGDDCPLVLGICRHAFHRH 107

QY 83 CLSRWL-----KTRQVCPLDNREW 101

Db 108 CIDKWIAAPTNPQFRAQCPLCRQDW 131

RESULT 11

S52511

hypothetical protein YDL008w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein D2900

C:Species: Saccharomyces cerevisiae

C:Date: 08-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004

C:Accession: S52511; S67540

R:Andre, B.; Visser, S.; Urrestarazu, L.

submitted to the EMBL Data Library, February 1995

A:Description: The sequence of a 42 kb segment located on the left arm of chromosome IV

A:Reference number: S52492

A:Accession: S52511

A:Molecule type: DNA

A:Residues: 1-165 <AND>

A:Cross-references: UNIPROT:Q12157; EMBL:Z48432; NID:G683669; PIDN:CAA88351.1; PID:G68366

A:Experimental source: strain S288C

R:Urrestarazu, L.A.; Andre, B.; Visser, S.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67535

A:Accession: S67540

A:Molecule type: DNA

A:Residues: 1-165 <URR>

A:Cross-references: EMBL:Z74056; NID:G1430968; PIDN:CAA98564.1; PID:e252972; PID:G143096

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:APC11

A:Cross-references: SGB:S0002166

A:Map position: 4L

Query Match 22.1%; Score 136; DB 2; Length 165;

Best Local Similarity 26.4%; Pred. No. 2.8e-07;

Matches 28; Conservative 19; Mismatches 33; Indels 26; Gaps 3;

QY 21 RFVKKNAVALWAWDIV-----VDNCAICRNHIMDLCTEQANQAS 62

Db 2 KVKINEVHSVFAMSHIPSTSDAANDPIGNDEDEVCGICRASVGTCPSCK----- 56

QY 63 ATSEECTVANGVCNHAHFHCISRWLKT---RQVCPLDNREWEFOK 105

Db 57 FPGDQCPLVIGLCHHNFHDHCIIYRWLDTPTSKGLCEWCRQTFLQK 102

RESULT 12

G90113

Rbpl protein [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C:Accession: G90113

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reil

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A99082; MUID:I1323671; PMID:I1323671

A:Accession: G90113

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-99 <DOU>

A:Cross-references: UNIPROT:Q9AVZ5; GB:AJ010592; NID:G12580758; PIDN:CAC27076.1; GSPDB:G

C:Genetics:

A:Map position: 2

A:Genome: nucleomorph

C:Keywords: nucleomorph

Query Match 21.5%; Score 132.5; DB 2; Length 99;

Best Local Similarity 33.3%; Pred. No. 4e-07;

Matches 24; Conservative 14; Mismatches 27; Indels 7; Gaps 1;

QY 40 DNCAICRNHIMDLCTEQ-----ANQASATSEECTVANGVCNHAHFHCISRWLKTRQ 92

Db 22 EKCAICRNLEDNADPFERVGKHSKNFLDQISKNCFLAYGRCGHSFLICICIENWILNK 81

QY 93 VCPLDNREWEFOK 104

Db 82 NCPLCSRIWVYE 93

RESULT 13

T20241

hypothetical protein H05L14.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T20241; T23049

R:Wilkinson, J.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19241

A:Accession: T20241

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-2160 <WIL>

A:Cross-references: UNIPROT:O17709; EMBL:Z75533; PIDN:CAA99823.1; GSPDB:GN00019; CESP:H0:

A:Experimental source: clone C54G4

R:Barlow, K.

submitted to the EMBL Data Library, October 1997

A:Reference number: Z19662

A:Accession: T23049

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-2160 <WIZ>

A:Cross-references: EMBL:Z99772; PIDN:CAB16922.1; GSPDB:GN00019; CESP:H05L14.2

A:Experimental source: clone H05L14

C:Genetics:

A:Gene: CESP:H05L14.2

A:Map position: 1

A:Introns: 49/3; 130/3; 190/3; 237/3; 290/1; 463/3; 507/1; 524/2; 545/1; 636/2; 663/2; 7

Query Match 18.9%; Score 116.5; DB 2; Length 2160;

Best Local Similarity 30.5%; Pred. No. 0.00044;

Matches 29; Conservative 11; Mismatches 38; Indels 17; Gaps 3;

QY 8 DTPSGTNGAGKRFVKKNAVALWAWDIV---DNCAICRNHIMDLCTEQANQASAT 64

Db 2068 DVPMPFSETVMKKMFQCSGYELDVVTEREEVVEEDGCLIC-TEITEEAVE----- 2117

QY 65 SEECTVANGVCNHAHFHCISRWLKTRQVCPLDNR 99

Db 2118 ----TVTCDTCTREYHYHCISRWLKINSVCPQCSR 2148

RESULT 14

T06680

hypothetical protein T17F15.100 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T06680

R:Queciet, F.; Choiane, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigue

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15793

A:Accession: T06680

A:Molecule type: DNA

A:Residues: 1-349 <QUE>

A:Cross-references: UNIPROT:Q9SU66; EMBL:AL049658; GSPDB:GN00061; ATSP:T17F15.100

A:Experimental source: cultivar Columbia; BAC clone T17F15

C:Genetics:

A:Gene: ATSP:T17F15.100

A:Map position: 3

A:Introns: 26/1; 81/2

F:203-254/Domain: RING finger homology <RRN>

Query Match 16.9%; Score 104; DB 2; Length 349;

```

Query Match      16.6%; Score 102.5; DB 2; Length 249;
Best Local Similarity 26.0%; Pred. No. 0.0016;
Matches 27; Conservative 22; Mismatches 38; Indels 17; Gaps 4;

QY      5 MDVPTPGTNSGAGKKRFEVKKNAVALW-----AWDIVVDNCATCRNHI-----MDL 52
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Ddb     125 IDMAITSSSSSRGRGWGEIKWIKTEFYKANKAAENLLID--SYCYNVLNVATTEENG 182
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      53 CIECQANQASATSEECTVAGVCNHAFFHFCISRWLTKTRQVCPL 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Ddb     183 CAICNEDYIEGSS---IVAKLPCDHEFHGDCINKWLQLNHMCPL 223
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: June 23, 2005, 15:30:28  
Job time : 40 secs

**THIS PAGE BLANK (USPTO)**

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

THIS PAGE BLANK (USPTO)

Best Available Copy



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2005, 15:18:51 ; Search time 176 Seconds  
(without alignments)  
314.230 Million cell updates/sec

Title: US-09-541-462B-2

Perfect score: 616

Sequence: 1 MAAMDVTPSGTNSGAGKK.....KTRQVCPLDNREWFQKYGH 108

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	616	100.0	1 RBX1_HUMAN	P62877 homo sapien
2	616	100.0	1 RBX1_MOUSE	P62878 mus musculus
3	616	100.0	2 Q641T4	Q641t4 xenopus lae
4	605	98.2	111 2 Q642J6	Q642j6 brachydanio
5	604	98.1	108 1 RBX1_SALSA	Q8GG64 salmo salar
6	539.5	87.6	174 2 Q709J7	Q709j7 anopheles g
7	534	86.7	108 1 RBXA_DROME	Q9W561 drosophila
8	523.5	85.0	114 2 Q8S3S0	Q8S3s0 oryza sativ
9	523.5	85.0	123 2 Q657Y3	Q657y3 oryza sativ
10	523	84.9	110 2 Q7Y042	Q7Y042 populus tom
11	514.5	83.5	118 1 RBXA_ARATH	Q940x7 arabidopsis
12	505	82.0	116 2 Q7SE71	Q7se71 neurospora
13	499	81.0	115 1 RBXB_ARATH	Q9M2b0 arabidopsis
14	493	80.0	110 1 RBX1_CABEL	Q23450 caenorhabdi
15	469	76.1	107 1 RBX1_SCHPO	O13959 schizosacch
16	400.5	65.0	122 1 RBXB_DROME	Q9Nhx0 drosophila
17	400	64.9	115 2 Q6BKZ3	Q6bkz3 debaryomyce
18	399	64.8	186 2 Q6CE99	Q6ce99 yarrowia li
19	385	62.5	121 1 RBX1_YEAST	Q82773 saccharomyc
20	384	62.3	108 2 Q75AV1	Q75av1 ashbya goss
21	383	62.2	107 2 Q7RQX0	Q7rqx0 plasmodium
22	378	61.4	107 2 Q77367	Q77367 plasmodium
23	348.5	56.6	110 2 Q6CP68	Q6cp68 kluyveromyc
24	308	50.0	92 2 Q8SWJ6	Q8swj6 encephalito
25	292	47.4	113 1 RBX2_HUMAN	Q9wtz1 mus musculus
26	290.5	47.2	113 1 RBX2_MOUSE	Q9wtz1 mus musculus
27	284.5	46.1	113 2 Q7P607	Q7p607 drosophila
28	284	46.2	109 2 Q7PNW0	Q7pnw0 anopheles g
29	281.5	45.7	96 2 Q6NU82	Q6nu82 xenopus lae
30	266	43.2	112 2 P91404	P91404 caenorhabdi
31	210	34.1	84 2 Q940X6	Q940x6 arabidopsis

32	209	33.9	84	1	AN11_MOUSE	Q9CPX9 mus musculus
33	208	33.8	84	1	AN11_HUMAN	Q9NYG5 homo sapien
34	202.5	32.9	91	2	Q7Q443	Q7q443 anopheles g
35	198	32.1	84	2	Q8H306	Q8h306 oryza sativ
36	197	32.0	85	2	Q7KA43	Q7ka43 drosophila
37	179	29.1	89	2	Q7QX96	Q7qx96 giardia lam
38	169.5	27.5	109	2	Q6C1E3	Q6c1e3 yarrowia li
39	164.5	26.7	94	1	Y1P3_SCHPO	Q9ut86 schizosacch
40	154	25.0	86	2	Q7RGA3	Q7rga3 plasmodium
41	151	24.5	89	2	Q6LFF9	Q6lff9 plasmodium
42	147	23.9	162	2	Q6CMU0	Q6cmu0 kluyveromyc
43	146.5	23.8	135	2	Q20052	Q20052 caenorhabdi
44	141	22.9	57	2	Q9M9L0	Q9m9l0 arabidopsis
45	140	22.7	149	2	Q6FQF6	Q6fqf6 candida gla

#### ALIGNMENTS

RESULT 1  
RBX1\_HUMAN  
ID RBX1\_HUMAN STANDARD; PRT; 108 AA.  
AC P62877; Q8N628; Q9DLS2; Q9WUK9; Q9Y254;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE RING-box protein 1 (Rbx1) (Regulator of cullins 1) (RING finger  
DE protein 75) (ZYP protein).  
GN Name=RBX1; Synonyms=RNF75, ROC1;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., INTERACTION WITH CULLINS, AND MUTAGENESIS OF  
RP CYS-53; CYS-56; CYS-75 AND HIS-77.  
RC TISSUE=Cervical carcinoma;  
RX MEDLINE=99247022; PubMed=10230407; DOI=10.1016/S1097-2765(00)80482-7;  
RA Ohta T., Michel J.J., Schottelius A.J., Xiong Y.;  
RT "ROC1, a homolog of Apc11, represents a family of cullin partners with  
RT an associated ubiquitin ligase activity.";  
RL Mol. Cell 3:535-541(1999).  
RN [2]  
RP SEQUENCE FROM N.A., AND IDENTIFICATION IN CBC(VHL) COMPLEX.  
RP MEDLINE=99234320; PubMed=10213691; DOI=10.1126/science.284.5414.657;  
RA Kamura T., Koeppe D.M., Conrad M.N., Skowyra D., Moreland R.J.,  
RA Iliopoulos O., Lane W.S., Kaelin W.G. Jr., Elledge S.J., Conaway R.C.,  
RA Harper J.W., Conaway J.W.;  
RT "Rbx1, a component of the VHL tumor suppressor complex and SCF  
RT ubiquitin ligase.";  
RL Science 284:657-661(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Collins J.E., Wright C.L., Edwards C.A., Davis M.P., Grinham J.A.,  
RA Cole C.G., Goward M.E., Aguado B., Mallia M., Mokrab Y., Huckle E.J.,  
RA Beare D.M., Dunham I.;  
RT "A genome annotation driven approach to cloning the human ORFeome.";  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20057165; PubMed=10591208; DOI=10.1038/990031;  
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,  
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
RA Dhani P.D., Dockree C., Dodsorth S.J., Durbin R.M., Ellington A.G.,  
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
RA Gilbert J.G.R., Goward M.E., Grahnam D.V., Griffiths M.N.D., Hall C.,  
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,

- RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,  
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
RA Vaudin M., Walli M., Wallis J.M., Whiteley M.N., Willey D.L.,  
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,  
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,  
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,  
RA Phan S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,  
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
RA Scheep P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,  
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston J.,  
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,  
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,  
RA Peyraud M., Kedra D., Serousi E., Fransson I., Tapia I., Bruder C.E.,  
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
RA Khan A.S., Lane L., Tiliahun Y., Wright H.,  
RA "The DNA sequence of human chromosome 22."  
RL Nature 402:489-495(1999).  
[5]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain, and Placenta;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Kzywinieki M.I., Skaleka U., Smalus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[6]  
RN SEQUENCE OF 14-108 FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=20106778; PubMed=10643962;  
RA Perin J.-P., Seddighi N., Charbonnier F., Goudou D., Belkadi L.,  
RA Rieger F., Alliel P.M.,  
RT "Genomic organization and expression of the ubiquitin-proteasome  
RT complex-associated protein Rbx1/ROCI/Hrt1."  
RL Cell. Mol. Biol. 45:1131-1137(1999).  
[7]  
RN SEQUENCE OF 92-105, INTERACTION WITH CUL1, AND IDENTIFICATION IN A  
RP COMPLEX WITH CUL1, SKP1 AND SKP2.  
RX TISSUE=Cervical carcinoma;  
RA MEDLINE=99247021; PubMed=10230406; DOI=10.1016/S1097-2765(00)80481-5;  
RA Tan P., Fuchs S.Y., Chen A., Wu K., Gomez C., Ronai Z., Pan Z.-Q.,  
RT "Recruitment of a ROCI-CUL1 ubiquitin ligase by Skp1 and HOS to  
RT catalyze the ubiquitination of I kappa B alpha."  
RL Mol. Cell 3:527-533(1999).
- [8]  
RN FUNCTION.  
RP MEDLINE=20047893; PubMed=10579999; DOI=10.1101/gad.13.22.2928;  
RA Kamura T., Conrad M.N., Yan Q., Conaway R.C., Conaway J.W.,  
RT "The Rbx1 subunit of SCF and VHL E3 ubiquitin ligase activates Rbx1  
RT modification of cullins Cdc53 and Cul2."  
RL Genes Dev. 13:2928-2933(1999).  
[9]  
RN FUNCTION, AND SUBCELLULAR LOCATION.  
RP MEDLINE=20481777; PubMed=11027288;  
RX DOI=10.1126/MCB.20.21.8185-8197.2000;  
RA Furukawa M., Zhang Y., McCarville J., Ohta T., Xiong Y.,  
RT "The CUL1 C-terminal sequence and ROCI are required for efficient  
RT nuclear accumulation, NEDD8 modification, and ubiquitin ligase  
RT activity of CUL1."  
RL Mol. Cell. Biol. 20:8185-8197(2000).  
[10]  
RN IDENTIFICATION IN E3 UBIQUITIN LIGASE COMPLEX WITH MUF1, AND  
RP IDENTIFICATION IN COMPLEXES WITH CUL5.  
RX MEDLINE=21380117; PubMed=11384984; DOI=10.1074/jbc.M103093200;  
RA Kamura T., Burian D., Yan Q., Schmidt S.L., Lane W.S., Querido E.,  
RA Branton P.E., Shilatifard A., Conaway R.C., Conaway J.W.,  
RT "Muf1, a novel elongin BC-interacting leucine-rich repeat protein that  
RT can assemble with CUL5 and Rbx1 to reconstitute a ubiquitin ligase."  
RL J. Biol. Chem. 276:29748-29753(2001).  
[11]  
RN IDENTIFICATION IN E3 UBIQUITIN LIGASE COMPLEX WITH MED8.  
RX MEDLINE=22155962; PubMed=12149480; DOI=10.1073/pnas.162424199;  
RA Brower C.S., Sato S., Tomomori-Sato C., Kamura T., Pause A.,  
RA Stearman R., Klausner R.D., Malik S., Lane W.S., Sorokina I.,  
RA Roeder R.G., Conaway J.W., Conaway R.C.,  
RT "Mammalian mediator subunit mMED8 is an elongin BC-interacting protein  
RT that can assemble with Cul2 and Rbx1 to reconstitute a ubiquitin  
RT ligase."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:10353-10358(2002).  
[12]  
RN IDENTIFICATION IN SCF-LIKE COMPLEX, AND INTERACTION WITH CUL7.  
RX MEDLINE=22388271; PubMed=12481031; DOI=10.1073/pnas.252646399;  
RA Dias D.C., Dolios G., Wang R., Pan Z.Q.,  
RT "CUL7: a DOC domain-containing cullin selectively binds Skp1.Fbx29 to  
RT form an SCF-like complex."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16601-16606(2002).  
[13]  
RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 19-108 IN COMPLEX WITH 17-776  
RP OF CUL1, AND X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) IN SCF COMPLEX WITH  
RP CUL1, SKP1 AND SKP2.  
RX MEDLINE=21959435; PubMed=11961546; DOI=10.1038/416703a;  
RA Zheng N., Schulman B.A., Song L., Miller J.J., Jeffrey P.D., Wang P.,  
RA Chu C., Koepf D.M., Ellledge S.J., Pagano M., Conaway R.C.,  
RA Conaway J.W., Harper J.W., Pavletich N.P.,  
RT "Structure of the Cull1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase  
RT complex."  
RL Nature 416:703-709(2002).  
CC -I- FUNCTION: Component of the SCF (SKP1-CUL1-F-box protein) and the  
CC CBC(VHL) (CUL2-elongin BC-VHL) E3 ubiquitin ligase complexes,  
CC which mediate the ubiquitination and subsequent proteasomal  
CC degradation of target proteins involved in cell cycle progression,  
CC signal transduction and transcription. Through the RING-type zinc  
CC finger, seems to recruit the E2 ubiquitination enzyme, like CDC34,  
CC to the complex and brings it into close proximity to the  
CC substrate. Probably also stimulates CDC34 autoubiquitination.  
CC Promotes the neddylation of CUL1 and probably CUL2.  
CC -I- PATHWAY: Ubiquitin conjugation; third step.  
CC -I- SUBUNIT: Part of a SCF complex consisting of CUL1, Rbx1, Skp1 and  
CC SKP2. Part of a SCF-like complex consisting of CUL1, Rbx1, Skp1  
CC and FBXW8. Part of CBC(VHL) complexes with elongin BC complex  
CC (TCEB1 and TCEB2), CUL2 or CUL5 and VHL. Part of multisubunit E3  
CC ubiquitin ligase complexes with elongin BC complex (TCEB1 and  
CC TCEB2), CUL2 and MED8; elongin BC complex (TCEB1 and TCEB2), CUL5  
CC and MUF1. Part of multisubunit complexes with elongin BC complex  
CC (TCEB1 and TCEB2), elongin A/TCEB3 or SOCS1 or WSB1 and CUL5.  
CC Interacts directly with CUL1 and probably also with CUL2, CUL3,  
CC CUL4A, CUL4B, CUL5 and CUL7. Probably interacts with CDC34.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.  
CC -!- TISSUE SPECIFICITY: Widely expressed.

Query Match 100.0%; Score 616; DB 1; Length 108;  
Best Local Similarity 100.0%; Pred. No. 6.6e-61;  
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMDVTPTSGTNSGAGKREVEVKNAVALAWMDIVVDNCAICRNHIMDLICISQANO 60  
DB 1 MAAAMDVTPTSGTNSGAGKREVEVKNAVALAWMDIVVDNCAICRNHIMDLICISQANO 60

QY 61 ASATSEECTVANGVNCNHAFPHFCISRWLKTRQVCPDLNREWEFKYKH 108  
DB 61 ASATSEECTVANGVNCNHAFPHFCISRWLKTRQVCPDLNREWEFKYKH 108

RESULT 2

RBX1 MOUSE  
ID RBX1 MOUSE STANDARD; PRT: 108 AA.  
AC P62878; Q8N628; Q9D1S2; Q9WUK9; Q9Y254;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DE 25-OCT-2004 (Rel. 45, Last annotation update)  
DE RING-box protein 1 (Rbx1).  
GN Name=Rbx1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND IDENTIFICATION IN CBC(VHL) COMPLEX.  
RX MEDLINE=99234320; PubMed=10213691; DOI=10.1126/science.284.5414.657;  
RA Kamura T., Koepf D.M., Conrad M.N., Skowry D., Moreland R.J.,  
RA Iliopoulos O., Lane W.S., Kaelin W.G. Jr., Ellledge S.J., Conaway R.C.,  
RA Harper J.W., Conaway J.W.;  
RA "Rbx1, a component of the VHL tumor suppressor complex and SCF  
RT ubiquitin ligase.";  
RL Science 284:657-661(1999).  
RN [2]

SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
EX MEDLINE=22354583; PubMed=124766851; DOI=10.1038/nature01266;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamana H., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reid J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wyszynski-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [3]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain, and Breast tumor;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards R., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]

RP IDENTIFICATION IN E3 UBIQUITIN LIGASE COMPLEX WITH MUF1, AND  
RP IDENTIFICATION IN COMPLEXES WITH CUL5.  
RX MEDLINE=21380117; PubMed=11384984; DOI=10.1074/jbc.M103093200;  
RA Kamura T., Burian D., Yan Q., Schmidt S.L., Lane W.S., Querido E.,  
RA Branton P.E., Shilatifard A., Conaway R.C., Conaway J.W.;  
RT "Muf1, a novel Elongin BC-interacting leucine-rich repeat protein that  
RL can assemble with Cul5 and Rbx1 to reconstitute a ubiquitin ligase.";  
RL J. Biol. Chem. 276:29748-29753(2001).  
RN [5]

TISSUE SPECIFICITY.  
RX MEDLINE=20106778; PubMed=10643962;  
RA Perin J.-P., Seddigi N., Charbonnier F., Goudou D., Belkadi L.,  
RA Rieger F., Alliel P.M.;  
RT "Genomic organization and expression of the ubiquitin-proteasome  
RL complex-associated protein Rbx1/ROCI/Hrt1.";  
RL Cell. Mol. Biol. 45:1131-1137(1999).  
CC -!- FUNCTION: Component of the SCF (E3-ubiquitin ligase complex) and the  
CC CBC(VHL) (CUL2-elongin BC-VHL) E3 ubiquitin ligase complexes,  
CC which mediate the ubiquitination and subsequent proteasomal  
CC degradation of target proteins involved in cell cycle progression,  
CC signal transduction and transcription. Through the RING-type zinc  
CC finger, seems to recruit the E2 ubiquitination enzyme, like CDC34,  
CC to the complex and brings it into close proximity to the  
CC substrate. Probably also stimulates CDC34 autoubiquitination.  
CC Promotes the neddylation of CUL1 and probably CUL2 (By  
CC similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Part of a SCF complex consisting of CUL1, Rbx1, SKP1 and  
CC SKP2. Part of a SCF-like complex consisting of CUL7, Rbx1, SKP1  
CC and FBXW8. Part of CBC(VHL) complexes with elongin BC complex  
CC (TCEB1 and TCEB2), CUL2 or CUL5 and VHL. Part of multisubunit E3  
CC ubiquitin ligase complexes with elongin BC complex (TCEB1 and  
CC TCEB2), CUL2 and MED8; elongin BC complex (TCEB1 and TCEB2), CUL5  
CC and MUF1. Part of multisubunit complexes with elongin BC complex  
CC (TCEB1 and TCEB2), elongin A/TCEB3 or SOCS1 or WSB1 and CUL5.  
CC Interacts directly with CUL1 and probably also with CUL2, CUL3,  
CC CUL4A, CUL4B, CUL5 and CUL7. Probably interacts with CDC34.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
CC -!- TISSUE SPECIFICITY: Widely expressed.  
CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
CC ubiquitin ligase activity. It coordinates an additional third zinc  
CC atom.  
CC -!- SIMILARITY: Belongs to the RING-box family.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC -----
CC EMBL; AF140599; AAD29716.1; -.
CC EMBL; AK003159; BAB22612.1; -.
CC EMBL; BC027396; -. NOT_ANNOTATED_CDS.
CC EMBL; BC056992; AAB56992.1; -.
CC DR MGD; MGI:1891829; Rbx1.
CC DR InterPro; IPR001841; Znf_ring.
CC DR Pfam; PF00097; zf-C3HC4_1.
CC DR PROSITE; PS00089; ZF_RING_2; 1.
CC KW Metal-binding; Nuclear protein; Ubl conjugation pathway; Zinc;
KW Zinc-finger.
FT ZN FING 53 98 RING-type.
FT METAL 42 42 Zinc 1 (By similarity).
FT METAL 45 45 Zinc 1 (By similarity).
FT METAL 53 53 Zinc 3 (By similarity).
FT METAL 56 56 Zinc 3 (By similarity).
FT METAL 68 68 Zinc 3 (By similarity).
FT METAL 75 75 Zinc 2 (By similarity).
FT METAL 77 77 Zinc 2 (By similarity).
FT METAL 80 80 Zinc 1 (By similarity).
FT METAL 82 82 Zinc 3 (By similarity).
FT METAL 83 83 Zinc 1 (By similarity).
FT METAL 94 94 Zinc 2 (By similarity).
FT METAL 97 97 Zinc 2 (By similarity).
FT CONFLICT 42 42 C -> F (in Ref. 2).
SQ SEQUENCE 108 AA; 12274 MW; 30FC5ADF66096C0E CRC64;

Query Match 100.0%; Score 616; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.6e-61;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAMVDVTPSGTNSGAGKKRFEVKWNAVALWANDIVVDNCAICRNHHIMDLICIEQANO 60
DB 1 MAAMVDVTPSGTNSGAGKKRFEVKWNAVALWANDIVVDNCAICRNHHIMDLICIEQANO 60

QY 61 ASATSECTVAMGVNCNHAFFHCISRWLKTRQVCLDNRWEFQKYGH 108
DB 61 ASATSECTVAMGVNCNHAFFHCISRWLKTRQVCLDNRWEFQKYGH 108

RESULT 3
Q641T4 PRELIMINARY; PRT; 108 AA.
AC Q641T4 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]_TaxID=8355;
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
```

```
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC082183; AAB82183.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 12274 MW; 30FC5ADF66096C0E CRC64;

Query Match 100.0%; Score 616; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 6.6e-61;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAMVDVTPSGTNSGAGKKRFEVKWNAVALWANDIVVDNCAICRNHHIMDLICIEQANO 60
DB 1 MAAMVDVTPSGTNSGAGKKRFEVKWNAVALWANDIVVDNCAICRNHHIMDLICIEQANO 60

QY 61 ASATSECTVAMGVNCNHAFFHCISRWLKTRQVCLDNRWEFQKYGH 108
DB 61 ASATSECTVAMGVNCNHAFFHCISRWLKTRQVCLDNRWEFQKYGH 108

RESULT 4
Q642J6 PRELIMINARY; PRT; 111 AA.
AC Q642J6
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RC TISSUE=liver;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
```

DR	InterPro; IPR001841; Znf_ring.
DR	Pfam; PF00097; zf-C3HC4; 1.
DR	PROSITE; PS50089; ZF_RING_2; 1.
DR	Metal-binding; Nuclear protein; Ubl conjugation pathway; Zinc;
KW	Zinc-finger.
FT	ZN_FING 53 98 RING-type.
FT	METAL 42 42 Zinc 1 (By similarity).
FT	METAL 45 45 Zinc 1 (By similarity).
FT	METAL 53 53 Zinc 3 (By similarity).
FT	METAL 56 56 Zinc 3 (By similarity).
FT	METAL 68 68 Zinc 3 (By similarity).
FT	METAL 75 75 Zinc 2 (By similarity).
FT	METAL 77 77 Zinc 2 (By similarity).
FT	METAL 80 80 Zinc 1 (By similarity).
FT	METAL 82 82 Zinc 3 (By similarity).
FT	METAL 83 83 Zinc 1 (By similarity).
FT	METAL 94 94 Zinc 2 (By similarity).
FT	METAL 97 97 Zinc 2 (By similarity).
SQ	SEQUENCE 108 AA; 12318 MW; 44C3EA712CEDC7BB CRC64;
Query Match 98.1%; Score 604; DB 1; Length 108;	
Best Local Similarity 98.1%; Pred. No. 1.4e-59;	
Matches 106; Conservative 0; Mismatches 2; Indels 0; Gaps 0	
QY	1 MAAMDVDPSTGNSGAGKKRFEVKKNAVALMAWDIVVDNCAICRNHIMDLCTECQAQ 60
DB	1 MAAMDVDPSTATNSGASKRRFEVKKNAVALMAWDIVVDNCAICRNHIMDLCTECQAQ 60
QY	61 ASATSEECTVANGVCNHAFHFHCISRWLKTROVCPDLDNRWEFKYGH 108
DB	61 ASATSEECTVANGVCNHAFHFHCISRWLKTROVCPDLDNRWEFKYGH 108
RESULT 6	
Q7Q9J7	
ID	PRELIMINARY; PRT; 174 AA.
AC	Q7Q9J7;
DT	01-WAR-2004 (TrEMBLrel. 26, Created)
DT	01-WAR-2004 (TrEMBLrel. 26, Last sequence update)
DT	01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	AgCP15018 (Fragment).
GN	Name=agCG46171; ORFNAMES=ENSGG00000007948;
OS	Anopheles gambiae str. PEST.
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OX	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
NCBI_TaxID=180454;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PEST;
RA	Anopheles Genome Sequencing Consortium;
RRL	Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC	-!- SIMILARITY: Contains 1 RING-type zinc finger
CC	-!- CAUTION: the sequence shown here is derived from an
CC	EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is preliminary data.
CC	EMBL; AAA001008900; EAA09407.1; --
DR	GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0008270; F;zinc ion binding; IEA.
DR	GO; GO:0016567; P:protein ubiquitination; IEA.
DR	InterPro; IPR001841; Znf ring.
PFam; PF00097; zf-C3HC4; 1.	
DR	PROSITE; PS50089; ZF_RING_2; 1.
KW	Metal-binding; Zinc; Zinc-finger.
NON_TER	1
SQ	SEQUENCE 174 AA; 20040 MW; E981973BC526D3A1 CRC64;
Query Match 87.6%; Score 539.5; DB 2; Length 174;	
Best Local Similarity 86.2%; Pred. No. 3.8e-52;	
Matches 94; Conservative 5; Mismatches 5; Indels 5; Gaps 1;	
QY	5 MDVD-----TPSGTNSGAGKKRFEVKKNAVALMAWDIVVDNCAICRNHIMDLCTECQAQ 59
DB	5 MDVD-----TPSGTNSGAGKKRFEVKKNAVALMAWDIVVDNCAICRNHIMDLCTECQAQ 59

Db 66 MDIDEEFEAPTSSSSRGRKRFEVKKNVNAWADIIVDNCAICRNHMDLCTECQAN 125

QY 60 QASATSECTVAGVNCNHAHFHCISRWLTKTQVCPDLNREWEFOKYGH 108

Db 126 QASATSECTVAGVNCNHAHFHCISRWLTKTQVCPDLNREWEFOKYGH 174

RESULT 7

REXA\_DROME STANDARD; PRT; 108 AA.

AC Q9W5E1; 077429;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 25-JAN-2005 (Rel. 46, Last annotation update)

DE RING-box protein 1A (Regulator of cullins 1a) (dbrxl).

GN Name=Roc1A; ORFNames=CG16982;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkely;

RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasearman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

RN [2]

RP GENOME REANNOTATION.

RX MEDLINE=22426069; PubMed=12537572;

RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

"Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review."

RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R;

RX MEDLINE=20196011; PubMed=10731137; DOI=10.1126/science.287.5461.2220;

RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,

RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,

RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,

RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,

RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,

RA Modolell J., Peter A., Schoettler P., Werner M., Mourikioti F.,

RA Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,

RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,

RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,

RA Glover D.M.;

RT "From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*."

RL Science 287:2220-2222(2000).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkely; TISSUE=Embryo;

RX MEDLINE=22426066; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,

RA George R.A., Guarin H., Kronmiller B., Pacleeb J.M., Park S., Wan K.H.,

RA Rubin G.M., Celniker S.E.;

RT "A *Drosophila* full-length cDNA resource."

RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).

RN [5]

RP FUNCTION, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND MUTAGENESIS OF ASN-59 AND CYS-68.

RX MEDLINE=22057626; PubMed=12062088; DOI=10.1016/S1534-5807(02)00164-8;

RA Nouredine M.A., Donaldson T.D., Thacker S.A., Duronio R.J.;

RT "Drosophila Roc1A encodes a RING-H2 protein with a unique function in processing the Hh signal transducer Ci by the SCF E3 ubiquitin ligase."

RT Dev. Cell 2:757-770(2002).

RN [6]

RP INTERACTIONS WITH LIN19 AND SLMB.

RX MEDLINE=21391618; PubMed=11500045; DOI=10.1006/bbrc.2001.5394;

RA Bocca S.N., Muzzopappa M., Silberstein S., Wappner P.;

RT "Occurrence of a putative SCF ubiquitin ligase complex in *Drosophila*."

RL Biochem. Biophys. Res. Commun. 286:357-364(2001).

RN [7]

RP REVIEW ON E3 UBIQUITIN LIGASE COMPLEXES.

RX MEDLINE=22736111; PubMed=12850443; DOI=10.1016/S0168-9525(03)00146-X;

RA Ou C.-Y., Pi H., Chien C.-T.;

RT "Control of protein degradation by E3 ubiquitin ligases in *Drosophila* eye development."

RL Trends Genet. 19:382-389(2003).

CC -!- FUNCTION: Component of the SCF (SKP1-CUL1-F-box protein) E3 ubiquitin ligase complex, which mediates the ubiquitination and subsequent proteasomal degradation of target proteins. Through the RING-type zinc finger, seems to recruit the E2 ubiquitination enzyme to the complex and brings it into close proximity to the substrate. Required for the specific SCF-dependent proteolysis of Ci, but not that of ARM, suggesting that it also participates in the selection of substrates inside the SCF complex.

CC -!- PATHWAY: Ubiquitin conjugation; third step.

CC -!- SUBUNIT: Part of a SCF complex consisting of SkpA (SKP1), Lin19 (CUL1), Roc1A and F-box protein Slmb. Interacts directly with Lin19 and Slmb.

CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.

CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in embryonic, larval and adult tissues.

CC -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.

CC -!- DOMAIN: The RING-type zinc finger domain is essential for ubiquitin ligase activity. It coordinates an additional third zinc atom (By similarity).

CC -!- SIMILARITY: Belongs to the RING-box family.

CC -!- SIMILARITY: Contains 1 RING-type zinc finger.

NCBI\_TaxID=39947;  
[1]  
SEQUENCE FROM N.A.  
Park Y.-J., Rostoks N., Ramakrishna W., SanMiguel P., Shiloff B.,  
Ma J., Jiang Z., Kleinhofs A., Bennetzen J.,  
Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A.  
Sasaki T., Matsumoto T., Katayose Y.,  
Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL: AF480496; AAL87158.1; -;  
DR EMBL: AP005844; BAD08109.1; -;  
DR Gramene; Q8S3S0; -;  
DR GO: GO:0000151; C:ubiquitin ligase complex; IEA.  
DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO: GO:0008270; F:zinc ion binding; IEA.  
DR GO: GO:0016567; P:protein ubiquitination; IEA.  
DR InterPro: IPR001841; Znf\_ring.  
DR Pfam: PF000097; zf-C3HC4.1.  
DR SMART; SM00184; RING.1.  
DR PROSITE; PS00089; 2F\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 114 AA; 12754 MW; 568C63E0D733DD5C CRC64;  
  
Query Match 85.0%; Score 523.5; DB 2; Length 114;  
Best Local Similarity 84.4%; Pred. NO. 1.5e-50;  
Matches 92; Conservative 6; Mismatches 10; Indels 1; Gaps 1;  
  
Qy 1 MAAMVDVTPSGTNSGAGK-KREVEKKNNAVALWANDIVVDNCAICRNHIMDLCEQAN 59  
Db 6 VAVVPPSIAGASSGAKGKREIFEIKKNVSLWAWDIVVDNCAICRNHIMDLCEQAN 65  
  
Qy 60 QASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFQYKH 108  
Db 66 QASATSEECTVANGVCNHAFFHCISRWLKTRQVCPLDNREWEFQYKH 114  
  
RESULT 9  
Q657Y3 PRELIMINARY; PRT; 123 AA.  
ID Q657Y3  
AC Q657Y3;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Putative ring-box protein.  
GN NamesP0005A05.23; Synonyms=P0482C06.1;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
ON NCBI\_TaxID=39947;  
OX [1]  
RX SEQUENCE FROM N.A.  
RP PubMed=12447438;  
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,  
Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,  
Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,  
Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,  
Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
Nagaoka H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,  
Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
Yano M., Jiang J., Gojobori T.;  
RT "The genome sequence and structure of rice chromosome 1.";  
RL Nature 420:312-316(2002).  
DR EMBL: AP002863; BAD44884.1; -;  
DR EMBL: AP002845; BAD44845.1; -;



```

SQ SEQUENCE 123 AA; 13625 MW; 38260146F7B4DBA7 CRC64;
Query Match 85.0%; Score 523.5; DB 2; Length 123;
Best Local Similarity 77.5%; Pred. No. 1.6e-50;
Matches 93; Conservative 5; Mismatches 5; Indels 17; Gaps 2;

QY 6 DVDTP-----SGTNSGAG-----KKRFEVKKNNAVALWAWDIIVDNCALICRNH 48
DB 4 DINAPPPAPAPAGAGEGSSSRAAGSRKPKRFEIKKNNAVALWAWDIIVDNCALICRNH 63
QY 49 INDLCIECQANQASATSECTVANGVNCNHFHFCISRLWLTQVCPDLNREWEFOKYGH 108
DB 64 INDLCIECQANQASATSECTVANGVNCNHFHFCISRLWLTQVCPDLNREWEFOKYGH 123

RESULT 10
ID Q7Y042 PRELIMINARY; PRT; 110 AA.
AC Q7Y042;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ring box protein.
OS Populus tomentosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
OX NCBI_TaxID=118781;
RN [1]
RP SEQUENCE FROM N.A.
RA Fan J.H.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AY302066; AAP57304.1; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF000097; zif-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
KW SEQUENCE 110 AA; 12591 MW; 0172C6CAE3FA2772 CRC64;

Query Match 84.9%; Score 523; DB 2; Length 110;
Best Local Similarity 84.5%; Pred. No. 1.6e-50;
Matches 93; Conservative 4; Mismatches 7; Indels 6; Gaps 2;

QY 5 MDVD---TPSG---TNSGACKRFEVKKNNAVALWAWDIIVDNCALICRNHMDLCIECQA 58
DB 1 MDTDTVTMPVAGEASSSSRRKPKRFEIKKNNAVALWAWDIIVDNCALICRNHMDLCIECQA 60
QY 59 NQASATSECTVANGVNCNHFHFCISRLWLTQVCPDLNREWEFOKYGH 108
DB 61 NQASATSECTVANGVNCNHFHFCISRLWLTQVCPDLNREWEFOKYGH 110

RESULT 11
ID RBXA ARATH STANDARD; PRT; 118 AA.
AC Q94077;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE RING-box protein 1a (RBX1a-At) (At-Rbx1;1) (RBX1-2).
GN Names=RBX1A; OrderedLocustNames=At5g20570; ORFNames=F7C8.160;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Ockresz L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=11130714; DOI=10.1038/35048507;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Du H., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McLaughlin B., Robben J., Grymonprez B., Zimmermann W.,
RA Rampegerger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
RA Weitzenegger T., Bothe G., Rose M., Hauf J., Berneiser S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana."
RL Nature 408:823-826(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Prover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172; DOI=10.1126/science.1088305;
RA Yamada K., Lim J., Dale J.C., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.-J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newman G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayaishiaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,
RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome."
RL Science 302:842-846(2003).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX PubMed=11910074; DOI=10.1126/science.1071006;
RA Seki M., Narusaka M., Kamiya A., Ishida J., Satou M., Sakurai T.,
RA Nakajima M., Enju A., Akiyama K., Oono Y., Muramatsu M.,
RA Hayaishiaki Y., Kawai J., Carninci P., Itoh M., Ishii Y., Arakawa T.,
RA Shibata K., Shinagawa A., Shinozaki K.;
RT "Functional annotation of a full-length Arabidopsis cDNA collection."
RL Science 296:141-145(2002).
RN [6]
RP FUNCTION, TISSUE SPECIFICITY, IDENTIFICATION IN A SCF COMPLEX WITH
RP CUL1 AND TIR1, AND INTERACTION WITH CUL1.
RX MEDLINE=22204438; PubMed=12215511;

```



RA Gray W.M., Hellmann H., Dharmasiri S., Estelle M.;  
 RT "Role of the Arabidopsis RING-H2 protein RBX1 in RUB modification and  
 RL SCF function."; Plant Cell 14:2137-2144(2002).  
 RN [7]  
 RP FUNCTION, TISSUE SPECIFICITY, IDENTIFICATION IN SCF COMPLEX, AND  
 RX INTERACTIONS WITH CUL1; CUL4; ASK1 AND ASK2  
 RX MEDLINE=22370998; PubMed=12381738; DOI=10.1074/jbc.M204254200;  
 RA Lechner E., Xie D., Grava S., Pigaglio E., Planchais S.,  
 RA Murray J.A.H., Parmentier Y., Mutterer J., Dubreucq B., Shen W.-H.,  
 RA Genschel P.;  
 RT "The AtRbx1 protein is part of plant SCF complexes, and its down-  
 RT regulation causes severe growth and developmental defects."; J.  
 RL J. Biol. Chem. 277:50069-50080(2002).  
 RN [8]  
 RP FUNCTION.  
 RX MEDLINE=22568282; PubMed=12682009; DOI=10.1093/emboj/cdg190;  
 RA Dharmasiri S., Dharmasiri N., Hellmann H., Estelle M.;  
 RT "The RUB/Nedd8 conjugation pathway is required for early development  
 RT in Arabidopsis."; EMBO J. 22:1762-1770(2003).  
 RL [9]  
 CC -!- FUNCTION: Component of the SCF (SKP1-CUL1-F-box protein) E3  
 CC ubiquitin ligase complex, which mediates the ubiquitination and  
 CC subsequent proteosomal degradation of target proteins. The SCF  
 CC complex plays a crucial role in regulating response to auxin and  
 CC is essential for growth and development. Through the RING-type  
 CC zinc finger, seems to recruit the E2 ubiquitination enzyme, to the  
 CC complex and brings it into close proximity to the substrate.  
 CC Promotes the neddylation of CUL1.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Part of SCF complexes, which consist of a SKP1-related  
 CC protein, a cullin, a RBX protein and a F-box protein. Part of a  
 CC SCF complex with ASK1 or ASK2 and CUL1. Part of a SCF complex with  
 CC CUL1 and TIR1. Interacts with CUL1 and CUL4.  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).  
 CC CUL1 and TIR1. Interacts with CUL1 and CUL4.  
 CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in shoot,  
 CC siliques, meristem, flowers, floral buds, open flowers, leaves,  
 CC stems, roots, germinant seeds and seedlings in dark. Expressed at a  
 CC higher level in tissues containing actively dividing cells.  
 CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity. It coordinates an additional third zinc  
 CC atom.  
 CC -!- SIMILARITY: Belongs to the RING-box family.  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; AF052401; AAL13435.1; -;  
 DR EMBL; AF296833; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AY086913; AAM64477.1; -;  
 DR EMBL; AY072430; AAL62422.1; -;  
 DR EMBL; AY114719; AAM48038.1; -;  
 DR EMBL; AK118181; BAC42804.1; -;  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 DR Metal-binding; Nuclear protein; Ubl conjugation pathway; Zinc;  
 KW Zinc-finger.  
 FT ZN FING 63 108 RING-type.  
 FT METAL 52 52 Zinc 1 (By similarity).  
 FT METAL 55 55 Zinc 1 (By similarity).  
 FT METAL 63 63 Zinc 3 (By similarity).  
 FT METAL 66 66 Zinc 3 (By similarity).  
 FT METAL 78 78 Zinc 3 (By similarity).  
 FT METAL 85 85 Zinc 2 (By similarity).  
 FT METAL 87 87 Zinc 2 (By similarity).  
 FT METAL 90 90 Zinc 1 (By similarity).

FT METAL 92 92 Zinc 3 (By similarity).  
 FT METAL 93 93 Zinc 1 (By similarity).  
 FT METAL 104 104 Zinc 2 (By similarity).  
 FT METAL 107 107 Zinc 2 (By similarity).  
 SQ SEQUENCE 118 AA; 13238 MW; 19947BF06F442A82 CRC64;  
 Query Match 83.5%; Score 514.5; DB 1; Length 118;  
 Best Local Similarity 78.6%; Pred. No. 1.5e-49;  
 Matches 92; Conservative 6; Mismatches 8; Indels 11; Gaps 2;  
 QY 3 AAMDVDT---PSG-----TNSGAGKREKVKKNVAVALWAWDIVVDNCAICRNHMD 51  
 DB 2 ATLSDVTMTIPAGEASSSVASSNNKAKREIKKWSAVALWAWDIVVDNCAICRNHMD 61  
 QY 52 LCIECQANQASATSECTVANGVNCNHFHFCISRWLTKTRQVCPLDNREWEFKYGH 108  
 DB 62 LCIECQANQASATSECTVANGVNCNHFHFCISRWLTKTRQVCPLDNREWEFKYGH 118  
 RESULT 12  
 Q7SBT1 PRELIMINARY; PRT; 116 AA.  
 AC Q7SBT1;  
 DT 01-MAR-2004 (TRENBLrel. 26, Created)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=NCU06224.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OR74A;  
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Inakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,  
 RA Kamal M., Kamysseilis M., Maucell E., Bielek C., Rudd S., Frishman D.,  
 RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,  
 RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
 RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0:0-0(2003).  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBSJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC EMBL; AABX01000163; EAA33839.1; -;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.  
 SQ SEQUENCE 116 AA; 13168 MW; BF1D46357C40D802 CRC64;  
 Query Match 82.0%; Score 505; DB 2; Length 116;  
 Best Local Similarity 94.4%; Pred. No. 1.7e-48;  
 Matches 85; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 18 GKKEFEVKKNVAVALWAWDIVVDNCAICRNHMDLCIECQANQASATSECTVANGVNCN 77  
 DB 26 GKPRFEVKKNVAVALWAWDIVVDNCAICRNHMDLCIECQANQASATSECTVANGVNCN 85

QY 78 AFHFHCISRWLKTQVCPDLNREWEFKYGH 107  
 DB 86 AFHFHCISRWLKTQVCPDLNREWEFKYGH 115

RESULT 13  
 RBXB ARATH  
 ID RBXB ARATH STANDARD; PRT; 115 AA.  
 AC Q9M2H0;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Putative RING-box protein 1b (RBX1b-At) (At-Rbx1;2) (RBX1-1).  
 GN Name=RBX1b; OrderedLocusNames=At3g42830; ORFNames=T21C14.50;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016720; PubMed=11130713; DOI=10.1038/35048706;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,  
 RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,  
 RA Dalseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,  
 RA De Simone V., Cholane N., Artiguenave F., Robert C., Brottier P.,  
 RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,  
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P.,  
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,  
 RA Conrad A., Horrischer K., Kauer G., Loehner T.-H., Nordsiek G.,  
 RA Reichelt J., Scharfe M., Schoen O., Barges M., Terol J., Climent J.,  
 RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,  
 RA Cooke R., Laude M., Berger-Llauro C., Purnelle B., Masuy D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casasuberta E.,  
 RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Walte A., Utterback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Matti R., Wu D., Peterson J., Van Aken S.,  
 RA Pai G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,  
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:820-822 (2000).  
 RN [2]  
 RN DISCUSSION OF SEQUENCE.  
 RP MEDLINE=22204438; PubMed=12215511;  
 RX Gray W.M., Hellmann H., Dharmasiri S., Estelle M.;  
 RA "Role of the Arabidopsis RING-H2 protein RBX1 in RUB modification and  
 SCF function.";  
 RT Plant Cell 14:2137-2144 (2002).  
 RN [3]  
 RN TISSUE SPECIFICITY.  
 RP MEDLINE=22370998; PubMed=12381738; DOI=10.1074/jbc.M204254200;  
 RX Lechner E., Xie D., Grava S., Pigaglio E., Planchais S., Shen W.-H.,  
 RA Murray J.A.H., Parmentier Y., Mutterer J., Dubreucq B., Shen W.-H.,  
 RA Genschik P.;  
 RT "The AtRbx1 protein is part of plant SCF complexes, and its down-  
 regulation causes severe growth and developmental defects.";  
 RL J. Biol. Chem. 277:50069-50080 (2002).  
 CC -1- FUNCTION: Potential component of the SCF (SKP1-CUL1-F-box protein)  
 CC E3 ubiquitin ligase complex, which mediates the ubiquitination and  
 CC subsequent proteasomal degradation of target proteins. The SCF  
 CC complex plays a crucial role in regulating response to auxin and

is essential for growth and development. Through the RING-type  
 zinc finger, seems to recruit the E2 ubiquitination enzyme, to the  
 complex and brings it into close proximity to the substrate (By  
 similarity).  
 -!- PATHWAY: Ubiquitin conjugation; third step.  
 -!- SUBUNIT: Potential part of SCF complexes, which consist of a SKP1-  
 related protein, a cullin, a RBX protein and a F-box protein.  
 -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).  
 -!- TISSUE SPECIFICITY: Not detected in floral buds, stems and roots.  
 -!- DOMAIN: The RING-type zinc finger domain is essential for  
 ubiquitin ligase activity. It coordinates an additional third zinc  
 atom.  
 -!- SIMILARITY: Belongs to the RBX1 family.  
 -!- SIMILARITY: Contains 1 RING-type zinc finger.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

EMBL; AL138639; CAB87200.1; --  
 F1R; T47341; T47341.  
 InterPro; IPR001841; Znf ring.  
 Pfam; PF00097; zf-C3HC4; 1.  
 PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Hypothetical protein; Metal-binding; Nuclear protein;  
 KW Ub1 conjugation pathway; Zinc; Zinc-finger.  
 FT ZING 60 107 RING-type.  
 FT METAL 49 49 Zinc 1 (By similarity).  
 FT METAL 52 52 Zinc 1 (By similarity).  
 FT METAL 60 60 Zinc 3 (By similarity).  
 FT METAL 63 63 Zinc 3 (By similarity).  
 FT METAL 75 75 Zinc 3 (By similarity).  
 FT METAL 82 82 Zinc 2 (By similarity).  
 FT METAL 84 84 Zinc 2 (By similarity).  
 FT METAL 87 87 Zinc 1 (By similarity).  
 FT METAL 89 89 Zinc 3 (By similarity).  
 FT METAL 90 90 Zinc 1 (By similarity).  
 FT METAL 101 101 Zinc 2 (By similarity).  
 FT METAL 104 104 Zinc 2 (By similarity).  
 SQ SEQUENCE 115 AA; 12999 MW; 1515E3E417DB1FAF CRC64;  
 Query Match 81.0%; Score 499; DB 1; Length 115;  
 Best Local Similarity 81.5%; Pred. No. 8.1e-48;  
 Matches 88; Conservative 5; Mismatches 13; Indels 2; Gaps 1;  
 QY 1 MAAAMDVDTFSGTNGAGKRFVKKNVAVALWMDIVVDNCAICRNHMDLCIECOANQ 60  
 DB 10 MGESSSISVPS--SSSKSKRFELKKWSAVALWMDIVVDNCAICRNHMDLCIECLANQ 67  
 QY 61 ASATSECTVAMGVCMHAFHFCISRWLKTQVCPDLNREWEFKYGH 108  
 DB 68 ASATSECTVAMGVCMHAFHFCISRWLKTQVCPDLNREWEFKYGH 115

RESULT 14  
 RBX1 CAEEL STANDARD; PRT; 110 AA.  
 AC Q23457; Q8WSQ1;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE RING-box protein 1 (Rbx1) (Ce-rbx-1).  
 GN Names=rbx-1; ORFNames=ZK287.5;  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RG "The C. elegans sequencing consortium;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE OF 2-110 FROM N.A.  
 RS Sasagawa Y., Urano T., Kohara Y., Takahashi H., Higashitani A.;  
 RA "Characterization of *rbx1* gene from *Caenorhabditis elegans*.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Component of the SCF (SKP1-CUL1-F-box protein) E3  
 CC ubiquitin ligase complex, which mediates the ubiquitination and  
 CC subsequent proteasomal degradation of target proteins. Through the  
 CC RING-type zinc finger, seems to recruit the E2 ubiquitination  
 CC enzyme to the complex and brings it into close proximity to the  
 CC substrate (By similarity).  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Part of SCF complexes, which consist of a SKP1 or a SKP1-  
 CC related protein, a cullin protein, and a F-box protein (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).  
 CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity. It coordinates an additional third zinc  
 CC atom.  
 CC -!- SIMILARITY: Belongs to the RING-box family.  
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-10 is the initiator.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Z70757; CA948401.1; -;  
 DR EMBL; AB077287; BAB83695.1; ALT INIT.  
 DR PIR; T27823; T27823.  
 DR WormBase; WBGene0004320; *rbx-1*.  
 DR WormPep; ZK287.5; CE06614.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 DR Metal-binding; Nuclear protein; Ubl conjugation pathway; Zinc;  
 KW Zinc-finger.  
 FT ZN\_FING 44 100 RING-type.  
 FT METAL 44 44 Zinc 1 (By similarity).  
 FT METAL 47 47 Zinc 1 (By similarity).  
 FT METAL 55 55 Zinc 3 (By similarity).  
 FT METAL 58 58 Zinc 3 (By similarity).  
 FT METAL 70 70 Zinc 3 (By similarity).  
 FT METAL 77 77 Zinc 2 (By similarity).  
 FT METAL 79 79 Zinc 2 (By similarity).  
 FT METAL 82 82 Zinc 3 (By similarity).  
 FT METAL 84 84 Zinc 3 (By similarity).  
 FT METAL 85 85 Zinc 1 (By similarity).  
 FT METAL 96 96 Zinc 2 (By similarity).  
 FT METAL 99 99 Zinc 2 (By similarity).  
 FT CONFLICT 2 3 AQ -> GP (in Ref. 2).  
 SQ SEQUENCE 110 AA; EPE50F7684B30A56 CRC64;  
 Query Match 80.0%; Score 493; DB 1; Length 110;  
 Best Local Similarity 79.3%; Pred. No. 3.6e-47;  
 Matches 88; Conservative 5; Mismatches 14; Indels 4; Gaps 2;  
 1 MAAAMD---VDPSTGNSGAGKRRFEVKKWNAVALWAWDVVDNCAICRNHIMDLICBQ 57  
 1 MAQASDSTAMEVEEATNTQTV-KRFEVKKWNAVALWAWDVVDNCAICRNHIMDLICBQ 59  
 58 ANQASATSEECTVAVGNCNHAFFHCISRWLKTQVCPDNRNEWFQKYGH 108

DB 60 ANQAGLKDCECTVAVGNCNHAFFHCISRWLKTQVCPDNRNEWFQKYGH 110  
 RESULT 15  
 REX1 SCHPO STANDARD; PRT; 107 AA.  
 ID REX1 SCHPO  
 AC O13959;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE RING-box protein p1p1 (Pop-interacting protein 1) (RING-box protein  
 DE 1).  
 GN Name=p1p1; ORFNames=SPAC23H4.18c;  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND INTERACTIONS  
 RP WITH POP1, POP2 AND PCUL1.  
 RX MEDLINE=22325332; PubMed=12167173; DOI=10.1186/1472-2091-3-22;  
 RA Seibert V., Prohl C., Schoutz I., Rhee E., Lopez R., Abderazzaq K.,  
 RA Zhou C., Wolf D.A.;  
 RA "Combinatorial diversity of fission yeast SCF ubiquitin ligases by  
 RT homo- and heterooligomeric assemblies of the F-box proteins Pop1 and  
 RT Pop2p.";  
 RL BMC Biochem. 3:22-22(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21859360; PubMed=11859360; DOI=10.1038/nature724;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sprouks J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: Component of E3 ubiquitin ligase SCF complexes, which  
 CC mediate the ubiquitination and subsequent proteasomal degradation  
 CC of target proteins. Seems to recruit the E2 ubiquitination enzyme,  
 CC like UBC3/CDC34, to the complex and brings it into close proximity  
 CC to the substrate.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Part of a SCF E3 ubiquitin ligase complex containing  
 CC psh1, p1p1, pcul1 and the F-box proteins pop1 and pop2. Instead of  
 CC the pop1/pop2 heterodimer the also homooligomers of pop1 or pop2 may  
 CC be present in the complex.  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
 CC -!- DOMAIN: The RING-type zinc finger domain is essential for  
 CC ubiquitin ligase activity. It coordinates an additional third zinc  
 CC atom.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2005, 00:17:37 ; Search time 2063 Seconds  
(without alignments)  
7680.492 Million cell updates/sec

Title: US-09-541-462B-1  
Perfect score: 327  
Sequence: 1 atggcgcagcagtgatgt.....tccaaaagatgggcactag 327

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pi.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_ats.\*  
12: gb\_av.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	327	100.0	327	9 AF142059	Homo sapi
2	327	100.0	433	6 CQ698451	Sequence
3	327	100.0	453	6 CQ690099	Sequence
4	327	100.0	467	6 CQ695007	Sequence
5	327	100.0	471	6 CQ712328	Sequence
6	327	100.0	472	6 CQ711142	Sequence
7	327	100.0	482	6 AX888031	Sequence
8	327	100.0	482	6 BD027641	Sequence
9	327	100.0	507	6 CQ729899	Sequence
10	327	100.0	508	6 BD271520	VonHippel
11	327	100.0	508	9 AF140598	Homo sapi
12	327	100.0	523	6 CQ701366	Sequence
13	327	100.0	535	9 BC001466	Homo sapi
14	327	100.0	535	9 CR456560	Homo sapi
15	327	100.0	554	9 BC017370	Homo sapi
16	320	97.9	503	6 CQ832496	Sequence
17	316	96.6	430	6 CQ709825	Sequence
18	308.8	94.4	4476	6 CQ493737	Sequence
19	308	94.2	497	9 HUMYQ60A05	Homo sapi

C 20	301.2	92.1	5347	6	AX281690	AX281690 Sequence
C 21	296.8	90.8	3484	9	HSTEST	X73608 H.sapiens m
22	290.2	88.7	504	6	BD271522	BD271522 VonHippel
23	290.2	88.7	504	10	AF140599	AF140599 Mus muscu
24	290.2	88.7	514	5	BC082183	BC082183 Xenopus l
25	290.2	88.7	531	10	BC051473	BC051473 Mus muscu
26	290.2	88.7	1612	10	BC056992	BC056992 Mus muscu
27	290.2	88.7	1616	10	BC027396	BC027396 Mus muscu
28	288.4	88.2	306	9	AY099360	AY099360 Homo sapi
29	287	87.8	153274	2	AY555280	AY555280 Mus muscu
30	285.4	87.3	172747	2	AC112970	AC112970 Mus muscu
31	281	85.9	468	6	CQ702730	CQ702730 Sequence
32	278.6	85.2	485	6	CQ707821	CQ707821 Sequence
33	273.8	83.7	475	6	CQ705678	CQ705678 Sequence
34	254	77.7	484	5	CR390140	CR390140 Gallus ga
35	253.6	77.6	3208	6	AR277693	AR277693 Sequence
36	253.6	77.6	3208	6	AR542209	AR542209 Sequence
37	253.6	77.6	3208	6	AX212267	AX212267 Sequence
38	249	76.1	300	6	CQ703590	CQ703590 Sequence
39	239	73.1	464	6	CQ713285	CQ713285 Sequence
40	237.4	72.6	536	5	BC081497	BC081497 Danio rer
41	235.8	72.1	1933	5	AY027936	AY027936 Salmo sal
42	231.6	70.8	104787	9	AC109638	AC109638 Homo sapi
43	230	70.3	433	11	G27926	G27926 human STS S
44	224.4	68.6	237747	2	AC123343	AC123343 Rattus no
45	224.4	68.6	275618	2	AC103413	AC103413 Rattus no

#### ALIGNMENTS

RESULT 1  
AF142059  
LOCUS AF142059 327 bp mRNA linear PRI 14-JUL-1999  
DEFINITION Homo sapiens RING finger protein (ROC1) mRNA, complete cds.  
ACCESSION AF142059  
VERSION AF142059.1 GI:4809215  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 327)  
AUTHORS Ohta, T., Michel, J.J., Schottelius, A.J. and Xiong, Y.  
TITLE ROC1, a homolog of APC11, represents a family of cullin partners with an associated ubiquitin ligase activity  
JOURNAL Mol. Cell 3 (4), 535-541 (1999)  
MEDLINE 99247022  
PubMed 10230407

REFERENCE 2 (bases 1 to 327)  
AUTHORS Ohta, T., Michel, J.J. and Xiong, Y.  
TITLE Direct Submission  
JOURNAL Submitted (10-APR-1999) Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Mason Farm Rd. and Manning Dr., Chapel Hill, NC 27599-7295, USA

#### FEATURES

source  
1..327  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..327  
/gene="ROC1"  
1..327  
/note="Cullin partnering protein with associated ubiquitin ligase activity"  
/codon\_start=1  
/product="RING finger protein"  
/protein\_id="AAD30146.1"  
/db\_xref="GI:4809215"  
/translation="MAAAMDVDTSGTNSGAGKKRFEVKKNAVALAWDVIWNCALCRNHMLDLCIECQANQASISECTVANGVNCNHAFPHCISRWLTKRQVCLDNRWEFKYGH"

## ORIGIN

Query Match 100.0%; Score 327; DB 9; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60

Db 1 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60

QY 61 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 120

Db 61 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 120

QY 121 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 180

Db 121 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 180

QY 181 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 240

Db 181 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 240

QY 241 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAGAG 300

Db 241 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAGAG 300

QY 301 TGGGAATTCAAAAGTATGGGCACTAG 327

Db 301 TGGGAATTCAAAAGTATGGGCACTAG 327

## RESULT 2

CQ698451  
LOCUS CQ698451 433 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 43377 from Patent WO02070737.  
ACCESSION CQ698451  
VERSION CQ698451.1 GI:42252856

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Liew, C.C., Marshall, W.E. and Zhang, H.

AUTHORS Compositions and methods relating to osteoarthritis

TITLE Patent: WO 02070737-A 43377 12-SEP-2002;

JOURNAL Chondrogene Inc. (CA)

## FEATURES

source  
1..433  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 327; DB 6; Length 433;  
Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60

Db 20 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 79

QY 61 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 120

Db 80 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 139

QY 121 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 180

Db 140 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 199

QY 181 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 240

Db 200 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 259

QY 241 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAG 300

Db 260 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAG 319

QY 301 TGGGAATTCAAAAGTATGGGCACTAG 327

Db 320 TGGGAATTCAAAAGTATGGGCACTAG 346

## RESULT 3

CQ690099  
LOCUS CQ690099 453 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 35025 from Patent WO02070737.  
ACCESSION CQ690099  
VERSION CQ690099.1 GI:42225927

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Liew, C.C., Marshall, W.E. and Zhang, H.

AUTHORS Compositions and methods relating to osteoarthritis

TITLE Patent: WO 02070737-A 35025 12-SEP-2002;

JOURNAL Chondrogene Inc. (CA)

## FEATURES

source  
1..453  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 327; DB 6; Length 453;  
Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60

Db 24 ATGCGCGCAGCGATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 83

QY 61 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 120

Db 84 CGCTTTGAAAGTGAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTTGTTGAT 143

QY 121 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 180

Db 144 AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGGCATAGAAATGTCAGCTAACCGAG 203

QY 181 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 240

Db 204 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGGAGTCTGTAAACCATGCTTTTCAC 263

QY 241 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAG 300

Db 264 TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACACACAGAG 323

QY 301 TGGGAATTCAAAAGTATGGGCACTAG 327

Db 324 TGGGAATTCAAAAGTATGGGCACTAG 350

## RESULT 4

CQ695007  
LOCUS CQ695007 467 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 39933 from Patent WO02070737.  
ACCESSION CQ695007  
VERSION CQ695007.1 GI:42240530

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE
AUTHORS      Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE        Compositions and methods relating to osteoarthritis
JOURNAL      Patent: WO 02070737-A 39933 12-SEP-2002;
              Chondrogene Inc. (CA)
FEATURES     Location/Qualifiers
source       1..467
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 327; DB 6; Length 467;
Best Local Similarity 100.0%; Pred. No. 1.4e-90;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 20 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 61 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 120
DB 80 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 139
QY 121 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 180
DB 140 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 199
QY 181 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 240
DB 200 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 259
QY 241 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 300
DB 260 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 319
QY 301 TGGGAATTCAAAAAGTATGGGCACCTAG 327
DB 320 TGGGAATTCAAAAAGTATGGGCACCTAG 346
RESULT 5
LOCUS      CQ712328              471 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 57254 from Patent WO02070737.
ACCESSION  CQ712328
VERSION     CQ712328.1 GI:42273185
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE       Compositions and methods relating to osteoarthritis
JOURNAL     Patent: WO 02070737-A 57254 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES     Location/Qualifiers
source       1..471
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 327; DB 6; Length 471;
Best Local Similarity 100.0%; Pred. No. 1.4e-90;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 17 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 76
QY 61 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 120
```

```
Db 77 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 136
QY 121 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 180
DB 137 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 196
QY 181 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 240
DB 197 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 256
QY 241 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 300
DB 257 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 316
QY 301 TGGGAATTCAAAAAGTATGGGCACCTAG 327
DB 317 TGGGAATTCAAAAAGTATGGGCACCTAG 343
RESULT 6
LOCUS      CQ711142              472 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION Sequence 56068 from Patent WO02070737.
ACCESSION  CQ711142
VERSION     CQ711142.1 GI:42271999
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Liew,C.C., Marshall,W.E. and Zhang,H.
TITLE       Compositions and methods relating to osteoarthritis
JOURNAL     Patent: WO 02070737-A 56068 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES     Location/Qualifiers
source       1..472
             /organism="Homo sapiens"
             /mol_type="unassigned DNA"
             /db_xref="taxon:9606"
ORIGIN
Query Match      100.0%; Score 327; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.4e-90;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 20 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 61 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 120
DB 80 CGCTTTGAAGTGAAAAAGTGGATGAGTCCCTCTGGGCTGGGATATTGTGTTGAT 139
QY 121 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 180
DB 140 AACTGTGCATCTCGAGGACCAACATTATGGATCTTTGCATAGATGTCAGCTAACCCAG 199
QY 181 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 240
DB 200 GCGTCCGCTACTTCCAGAGAGTGTACTGTCGATGGGAGTCTGTAACCATGCTTTTCAC 259
QY 241 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 300
DB 260 TTCACCTGCATCTCGCTGGTCCAAACACAGCAGGTGTGTCCATTGGACAACAGAGAG 319
QY 301 TGGGAATTCAAAAAGTATGGGCACCTAG 327
DB 320 TGGGAATTCAAAAAGTATGGGCACCTAG 346
RESULT 7
```

AX888031	AX888031	482 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	Sequence 3894 from Patent EP1033401.				
DEFINITION	AX888031				
ACCESSION	AX888031.1	GI:40046785			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.			
AUTHORS	Expressed sequence tags and encoded human proteins				
TITLE	Patent: EP 1033401-A 3894 06-SEP-2000;				
JOURNAL	Genset (FR)				
FEATURES	Location/Qualifiers				
source	1..482				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	29..355				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAE93956.1"				
	/db_xref="GI:40046786"				
	/translation="MAAMDVDTPSGTNSGAGKRFKVKWNAVALWAWDIIVVDNCAI				
	CRNHIMDLCTIECQANQASATSECTVANGVCHAFPHFICISRLWKTRQVCPDLNREWE				
	FKYXGH"				
ORIGIN					
Query Match	100.0%; Score 327; DB 6; Length 482;				
Best Local Similarity	100.0%; Pred. No. 1.4e-90;				
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	ATGCGCGCAGCGATGGATGTGGATACCCGAGCGGCGACCAACAGCGCGCGGCGGCAAGAAG	60		
Db	29	ATGCGCGCAGCGATGGATGTGGATACCCGAGCGGCGACCAACAGCGCGCGGCGGCAAGAAG	88		
Qy	61	CGCTTTGAAGTGA AAAAGTGAATGCAGTAGCCCTCTGGCGCTGGGATATGTGGTTGAT	120		
Db	89	CGCTTTGAAGTGA AAAAGTGAATGCAGTAGCCCTCTGGCGCTGGGATATGTGGTTGAT	148		
Qy	121	AACGTGTGCCATCTGCAGGAACCAATTTATGGATCTTTGCATAGAAATGCTCAAGCTAACCCAG	180		
Db	149	AACGTGTGCCATCTGCAGGAACCAATTTATGGATCTTTGCATAGAAATGCTCAAGCTAACCCAG	208		
Qy	181	CGCTCCCGTACTTCAGAAAGTAGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTTCAC	240		
Db	209	CGCTCCCGTACTTCAGAAAGTAGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTTCAC	268		
Qy	241	TTCCACTGCGATCTCTCGCTGGCTCAAAACAGCAGCAGGTGTCCATTTGGACACACAGAG	300		
Db	269	TTCCACTGCGATCTCTCGCTGGCTCAAAACAGCAGCAGGTGTCCATTTGGACACACAGAG	328		
Qy	301	TGGGAATTC AAAAGTATGGGCACTAG 327			
Db	329	TGGGAATTC AAAAGTATGGGCACTAG 355			
RESULT 8					
BD027641	BD027641	482 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	Sequence tag and encoded human protein.				
DEFINITION	BD027641				
ACCESSION	BD027641				
VERSION	BD027641.1	GI:22569383			
KEYWORDS	JP 2001269182-A/3887.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	(bases 1 to 482)			
AUTHORS	Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.				
TITLE	Sequence tag and encoded human protein				



```
/mol_type="unassigned DNA"  
/db xref="taxon:9606"
```

## ORIGIN

Query Match	100.0%	Score 327;	DB 6;	Length 507;
Best Local Similarity	100.0%	Pred. No. 1.4e-90;		
Matches 327; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ATGCGCGCAGGATGTGGATATACCCGAGCGCGCACCAACGCGCGGGCAAGAAG	60	
DB	7	ATGCGCGCAGGATGTGGATATACCCGAGCGCGCACCAACGCGCGGGCAAGAAG	66	
QY	61	CGCTTTGAAGTGA AAAAGTGGAAATGCAAGTAGCCCTCTGGGCTCGGGATATTGTGGTTGAT	120	
DB	67	CGCTTTGAAGTGA AAAAGTGGAAATGCAAGTAGCCCTCTGGGCTCGGGATATTGTGGTTGAT	126	
QY	121	AACGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGAAATGTCAAGCTTAACAG	180	
DB	127	AACGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGAAATGTCAAGCTTAACAG	186	
QY	181	GGGTCCGCTACTTTCAGAAAGTGTACTGTGCGCATGGGGAGTCTCTAACCATGCTTTTCAC	240	
DB	187	GGGTCCGCTACTTTCAGAAAGTGTACTGTGCGCATGGGGAGTCTCTAACCATGCTTTTCAC	246	
QY	241	TTCCATCTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAAACAGAGAG	300	
DB	247	TTCCATCTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAAACAGAGAG	306	
QY	301	TGGGAATTC AAAAGTATGGGCATAG	327	
DB	307	TGGGAATTC AAAAGTATGGGCATAG	333	

RESULT	PC
BD271520	508 bp DNA linear PAT 17-JUL-2003
LOCUS	VonHippel-Lindau tumor suppressor complex and novel component of
DEFINITION	SCF ubiquitin ligase.
ACCESSION	BD271520
VERSION	BD271520.1 GI:33081288
KEYWORDS	JP 2002541775-A/1.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Rutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	Conaway J.W., Conaway,R.C. and Kamura,T. VonHippel-Lindau tumor suppressor complex and novel component of SCF ubiquitin ligase
JOURNAL	Patent: JP 2002541775-A 1 10-DEC-2002; OKLAHOMA MEDICAL RESEARCH FOUNDATION
COMMENT	OS Homo sapiens (human) PN JP 2002541775-A/1 PD 10-DEC-2002 PF 25-FEB-2000 JP 2000601023 PR 26-FEB-1999 US 60/121787 PI JOAN W CONAWAY, RONALD C CONAWAY, TAKUMI KAMURA PC C12N15/09,A61K38/00,A61K38/53,A61K45/00,A61P35/00,C07K14/47, PC C12N1/15, PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12P21/02,G01N33/15,G01N33/PC 50, PC G01N33/68 CC VonHippel-Lindau tumor suppressor complex and novel component of SCF CC ubiquitin ligase FH Key Location/Qualifiers FT source 1..508 .PT /organism='Homo sapiens (human)'. Location/Qualifiers 1..508 /organism='Homo sapiens' /mol type='genomic DNA'
FEATURES	source

/db xref="taxon:9606"

## ORIGIN

Query Match	100.0%;	Score 327;	DB 6;	Length 508;
Best Local Similarity	100.0%;	Pred. No. 1.4e-90;		
Matches 327; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGCGCAGCGATGTGGATACCCGAGCGGCACCAACAGCGCGCGGCAACAAG	60	
Db	7	ATGCGGCAGCGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAACAAG	66	
Qy	61	CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATGTGGTTGAT	120	
Db	67	CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATGTGGTTGAT	126	
Qy	121	AACGTGGCCATCTGCAGGAACCAATATGGATCTTTGCATAGAAATGCTAAGCTAACCAAG	180	
Db	127	AACGTGGCCATCTGCAGGAACCAATATGGATCTTTGCATAGAAATGCTAAGCTAACCAAG	186	
Qy	181	GCGTCCGCTACTTTCAGAAGAGTCTACTGTCGCATGGGGAGTCTGTAACTGCTTTTTCAC	240	
Db	187	GCGTCCGCTACTTTCAGAAGAGTCTACTGTCGCATGGGGAGTCTGTAACTGCTTTTTCAC	246	
Qy	241	TTCCACTGCAATCTCTCGCTGGCTCAAAACACGACGAGTGTGTCCATTGGACAAACAGAG	300	
Db	247	TTCCACTGCAATCTCTCGCTGGCTCAAAACACGACGAGTGTGTCCATTGGACAAACAGAG	306	
Qy	301	TGGGAATTCAAAAGATATGGGCACTAG	327	
Db	307	TGGGAATTCAAAAGATATGGGCACTAG	333	

RESULT 11

```

AF140598          508 bp      mRNA      linear      PRI 11-MAY-1999
LOCUS              Homo sapiens ring-box protein 1 (RBX1) mRNA, complete cds.
DEFINITION
ACCESSION          AF140598
VERSION            AF140598.1 GI:4769003
KEYWORDS
SOURCE
ORGANISM            Homo sapiens (human)
                    Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1 (bases 1 to 508)
AUTHORS             Kamura,T., Koepf,D.M., Conrad,M.N., Skowrya,D., Moreland,R.J.,
                    Iliopoulos,O., Lane,W.S., Kaelin,W.G. Jr., Elledge,S.J.,
                    Conaway,R.C., Harper,J.W. and Conaway,J.W.
                    RBX1, a component of the VHL tumor suppressor complex and SCF
                    ubiquitin ligase
                    Science 284 (5414), 657-661 (1999)
JOURNAL             Science 284 (5414), 657-661 (1999)
MEDLINE             99234320
PUBMED              10213691
REFERENCE            2 (bases 1 to 508)
AUTHORS             Kamura,T., Lane,W.S., Conaway,R.C. and Conaway,J.W.
TITLE              Direct Submission
JOURNAL             Submitted (05-APR-1999) Prog. Molec. Cell Biol., HHMI, Oklahoma
                    Med. Res. Fdn., 825 NE 13th St, Oklahoma City, OK 73104, USA
                    Location/Qualifiers
FEATURES             source
                    1..508
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    gene
                    1..508
                    /gene="RBX1"
                    CDS
                    7..333
                    /gene="RBX1"
                    /note="ring finger-like protein; component of VHL tumor
                    suppressor complex and SCF ubiquitin ligase"
                    /codon_start=1
                    /product="ring-box protein 1"
                    /protein_id="AAD29715.1"
                    /db_xref="GI:4769004"
                    /translation="MAAAMDVTPSGTNSGAGKRFEVKKNAVALAWMDIVVDNCAI

```





ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 554)	
	Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauser, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, W.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
MEDLINE	22388257	
PUBMED	12477932	
REFERENCE	2 (bases 1 to 554)	
AUTHORS	Strausberg, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
COMMENT	Contact: MGC help desk Email: <a href="mailto:gcgaps@mail.nih.gov">gcgaps@mail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Institute for Systems Biology <a href="http://www.systemsbio.org">http://www.systemsbio.org</a> contact: <a href="mailto:amadansystemsbiology.org">amadansystemsbiology.org</a> Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting	
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 19 Row: c Column: 17 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22091459.</p> <p>Location/Qualifiers</p> <p>1..554</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:465797"</p> <p>/tissue_type="Brain, glioblastoma"</p> <p>/clone_lib="NIH_MGC_57"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pDNR-LIB"</p> <p>&lt;1..554</p> <p>/gene="RBX1"</p> <p>/note="synonyms: ROC1, BA554C12.1, MGC13357, MGC1481, RNF75"</p> <p>/db_xref="LocusID:9978"</p> <p>/db_xref="MIM:603814"</p> <p>&lt;1..349</p> <p>/gene="RBX1"</p> <p>/codon_start=2</p> <p>/product="RBX1 protein"</p> <p>/protein_id="AAH17370.2"</p> <p>/db_xref="GI:32425477"</p>	
source	<p>misc_feature</p> <p>80..346</p> <p>/db_xref="LocusID:9978"</p> <p>/db_xref="MIM:603814"</p> <p>/translation="GRPCVSKMAAADVDTPSGTNSGAGKRFVKKNVAVALWAWDI VVDNCAICRNHIMDLICISQANQASATSEECTVANGVCNHAHFHFCISRWLKTQVCP LDRNWEFKYGH"</p> <p>/gene="RBX1"</p> <p>/note="APC11; Region: COG5194, APC11, Component of SCF ubiquitin ligase and anaphase-promoting complex [Posttranslational modification, protein turnover, chaperones / Cell division and chromosome partitioning]"</p> <p>/db_xref="CDD:COG5194"</p>	
gene	<p>Query Match 100.0%; Score 327; DB 9; Length 554;</p> <p>Best Local Similarity 100.0%; Pred. No. 1.4e-90;</p> <p>Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>	
CDS	<p>1 ATGGCGGACGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60</p> <p>23 ATGGCGGACGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 82</p> <p>61 CGCTTTGAAGTGAAAAGTGGAAATGCAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT 120</p> <p>83 CGCTTTGAAGTGAAAAGTGGAAATGCAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT 142</p> <p>121 AACTGTGCCATCTGCAGGAACCAACATTTATGGATCTTTGCATAGAAATGTCAAGCTAACCCAG 180</p> <p>143 AACTGTGCCATCTGCAGGAACCAACATTTATGGATCTTTGCATAGAAATGTCAAGCTAACCCAG 202</p> <p>181 GGTCCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 240</p> <p>203 GGTCCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 262</p> <p>241 TTCCACTGCATCTCTCGCTGGCTCAAAAACAGACAGGTGTGTCCATTGGACACACAGAGAG 300</p> <p>263 TTCCACTGCATCTCTCGCTGGCTCAAAAACAGACAGGTGTGTCCATTGGACACACAGAGAG 322</p> <p>301 TGGGAATTCCAAAGTATGGGCACTAG 327</p> <p>323 TGGGAATTCCAAAGTATGGGCACTAG 349</p>	
	<p>Search completed: June 24, 2005, 05:57:24</p> <p>Job time : 2069 secs</p>	
	<p>ORIGIN</p> <p>Query Match 100.0%; Score 327; DB 9; Length 554;</p> <p>Best Local Similarity 100.0%; Pred. No. 1.4e-90;</p> <p>Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>	

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 18:42:59 ; Search time 359 Seconds  
(without alignments)  
5392.076 Million cell updates/sec

Title: US-09-541-462B-1

Perfect score: 327

Sequence: 1 atggggcagcgatgatgt.....tccaaaagatgggcactag 327

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	327	3	AAA96882 Nucleotid
2	327	100.0	482	3	AAC03896 Human sec
3	327	100.0	506	12	ADQ87496 Human tum
4	327	100.0	506	12	ADQ87156 Human tum
5	327	100.0	506	13	ADQ84881 Human tum
6	327	100.0	508	3	AAA74978 DNA encod
7	327	100.0	508	13	ACN40951 Tumour-as
8	325	99.4	476	9	ACH29979 Human tes
9	321.8	98.4	3726	13	ADS05913 Human tes
10	321.8	98.4	4543	11	ACN90161 Breast ca
11	320	97.9	503	12	ADQ92179 Human aut
12	308.8	94.4	4476	5	ABV25615 Human pro
13	302.4	92.5	380	8	ABX39512 Bovine ES
14	301.2	92.1	5347	6	AAS94844 Human DNA
15	296.8	90.8	3484	12	ADF42703 Human Tes
16	296.8	90.8	3484	12	ADQ18564 Human sof
17	296.8	90.8	3484	13	ADQ80865 Human SPA
18	296.8	90.8	5111	10	ADB75558 Prostate
19	296.8	90.8	5371	12	ADQ23032 Human sof
20	290.2	88.7	504	3	AAA74980 DNA encod

21	282.8	86.5	539	5	AAH97860	Aah97860 Murine 7-
22	282.8	86.5	586	5	AAH97862	Aah97862 Murine 7-
23	253.6	77.6	3208	4	AAD12859	Aad12859 Human nov
24	246.4	75.4	557	5	AAH97861	Aah97861 Murine 7-
25	219.2	67.0	485	3	AAA43288	Aaa43288 Xenopus s
26	209	63.9	509	5	AAS86845	Aas86845 DNA encod
27	205.8	62.9	439	9	ACH27559	Ach27559 Human adu
28	196.6	60.1	692	4	ABL22527	Ab122527 Drosophil
29	191	58.4	1050	10	ADD17060	Add17060 DNA (SeqI
30	191	58.4	1050	10	ADD17933	Add17933 DNA (SeqI
31	191	58.4	1050	10	ADK51138	Adk51138 Plant DNA
32	191	58.4	1050	10	ADK56974	Adk56974 Plant DNA
33	190	58.1	608	13	ADR59906	Adr59906 Cotton CD
34	188.4	57.6	544	13	ADR59907	Adr59907 Cotton CD
35	186.8	57.1	553	13	ACN55764	Acn55764 Cotton an
36	184.6	56.5	731	10	ADK56881	Adk56881 Plant DNA
37	184.4	56.4	617	10	ADK56882	Adk56882 Plant DNA
38	183.8	56.2	175561	8	AAD55694	Aad55694 Human THB
39	183.8	56.2	175561	12	ADL08129	Adl08129 Human Gen
40	182.8	55.9	527	10	ADK56883	Adk56883 Plant DNA
41	182	55.7	531	10	ADK56888	Adk56888 Plant DNA
42	182	55.7	586	10	ADK56887	Adk56887 Plant DNA
43	180.8	53.3	830	12	ADP22559	Adp22559 Sea-squir
44	175.6	53.7	691	10	ADK56880	Adk56880 Plant DNA
45	174	53.2	254	2	AAT26036	Aat26036 Human gen

## ALIGNMENTS

## RESULT 1

AAA96882  
ID AAA96882 standard; DNA; 327 BP.

XX AAA96882;

DT 19-FEB-2001 (first entry)

XX Nucleotide sequence of human ring finger protein ROC1.

XX ROC1; ROC2; cullin; ring finger protein; APC11; APC complex; SCF pathway;  
XX cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;  
XX tumour; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
XX CDS 1..327

XX /\*tag= a  
XX /product= "ROC1"

XX WO200058472-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008592.

XX 31-MAR-1999; 99US-0127261P.

XX 22-NOV-1999; 99US-0166927P.

XX (UYNC-) UNIV NORTH CAROLINA.

XX Xiong Y, Ohta T;

XX WPI: 2000-647235/62.

XX P-PSDB; AAB19160.

XX Novel nucleic acid encoding cullin regulating ring finger proteins,  
XX termed as ROC proteins similar to anaphase-promoting complex 11, for  
XX therapeutic and diagnostic use.

XX Claim 1; Fig 2A; 83pp; English.

XX

CC The present sequence encodes a human ROC1 ring finger protein. The  
 CC specification also describes human ROC2. ROC1 and ROC2 are similar to  
 CC APC11, a subunit of the APC complex. The proteins stimulate cullin  
 CC dependent ubiquitin ligase activity. ROC1 functions in vivo as an  
 CC essential regulator of CDK inhibitor Sic1 degradation by the SCF  
 CC (undefined) pathway. ROC proteins are useful for screening bioactive  
 CC agents that interfere with the binding of ROC proteins with cullin  
 CC proteins. Pharmaceutical formulations comprising ROC proteins are useful  
 CC for diagnostic and therapeutic purposes, preferably for diagnosing and  
 CC treating tumours  
 XX  
 SQ Sequence 327 BP; 85 A; 75 C; 94 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 327; DB 3; Length 327;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-101;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGCGCGCAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGCAAGAAG 60  
 DB 1 ATGCGCGCAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGCAAGAAG 60  
 QY 61 CGCTTTGAAAGTGAAGTGAATGCAATATGATGATCTTGGCCCTGGGATATTGTTGAT 120  
 DB 61 CGCTTTGAAAGTGAAGTGAATGCAATATGATGATCTTGGCCCTGGGATATTGTTGAT 120  
 QY 121 AACTGTGCCATCTGCAGGAAACCAATATGATGATCTTGGCATAGATGTCAAGCTAACCCAG 180  
 DB 121 AACTGTGCCATCTGCAGGAAACCAATATGATGATCTTGGCATAGATGTCAAGCTAACCCAG 180  
 QY 181 GCCTTCGCTACTTCAAGAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTTCAC 240  
 DB 181 GCCTTCGCTACTTCAAGAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTTCAC 240  
 QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACGACAGGTGTGTCATTGGACACACAGAG 300  
 DB 241 TTCCACTGCATCTCGCTGGCTCAAAACACGACAGGTGTGTCATTGGACACACAGAG 300  
 QY 301 TGGGAATTCAAAAGTATGGGCACTAG 327  
 DB 301 TGGGAATTCAAAAGTATGGGCACTAG 327

RESULT 2  
 AAC03896  
 ID AAC03896 standard; cDNA; 482 BP.  
 XX AAC03896;  
 XX  
 XX 06-OCT-2000 (first entry)  
 XX Human secreted protein 5' EST, SEQ ID NO: 3894.  
 XX  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX Homo sapiens.  
 XX EP1033401-A2.  
 XX 06-SEP-2000.  
 XX 21-FEB-2000; 2000EP-00200610.  
 XX 26-FEB-1999; 99US-0122487P.  
 XX (GENSET).  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI: 2000-500381/45.  
 DR P-PSDB; AAG03890.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 XX Claim 1; SEQ ID NO 3894; 71pp + Sequence Listing; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond  
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors  
 XX  
 SQ Sequence 482 BP; 116 A; 103 C; 121 G; 140 T; 0 U; 2 Other;

Query Match 100.0%; Score 327; DB 3; Length 482;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-101;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGCGCGCAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGCAAGAAG 60  
 DB 29 ATGCGCGCAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGCAAGAAG 88  
 QY 61 CGCTTTGAAAGTGAAGTGAATGCAATATGATGATCTTGGCCCTGGGATATTGTTGAT 120  
 DB 89 CGCTTTGAAAGTGAAGTGAATGCAATATGATGATCTTGGCCCTGGGATATTGTTGAT 148  
 QY 121 AACTGTGCCATCTGCAGGAAACCAATATGATGATCTTGGCATAGATGTCAAGCTAACCCAG 180  
 DB 149 AACTGTGCCATCTGCAGGAAACCAATATGATGATCTTGGCATAGATGTCAAGCTAACCCAG 208  
 QY 181 GCCTTCGCTACTTCAAGAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTTCAC 240  
 DB 209 GCCTTCGCTACTTCAAGAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTTCAC 268  
 QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACGACAGGTGTGTCATTGGACACACAGAG 300  
 DB 269 TTCCACTGCATCTCGCTGGCTCAAAACACGACAGGTGTGTCATTGGACACACAGAG 328  
 QY 301 TGGGAATTCAAAAGTATGGGCACTAG 327  
 DB 329 TGGGAATTCAAAAGTATGGGCACTAG 355

RESULT 3  
 ADQ87496  
 ID ADQ87496 standard; cDNA; 506 BP.  
 XX ADQ87496;  
 XX  
 XX 07-OCT-2004 (first entry)  
 XX Human tumour-associated antigenic target (TAT) cDNA sequence #4374.  
 DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
 KW cancer; cell proliferative disorder; gene; ss.  
 XX Homo sapiens.  
 XX WO2004060270-A2.  
 PN 22-JUL-2004.  
 PD 15-OCT-2003; 2003WO-US029126.  
 PF 19-OCT-2002; 2002US-0418988P.  
 XX

XX (GETH ) GENENTECH INC.  
PA (WUTD/) WU T D.  
PA (ZHOU/) ZHOU Y.  
XX  
PI Wu TD, Zhou Y;  
XX WPI; 2004-534300/51.  
DR  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.  
XX  
XX Claim 1; SEQ ID NO 4374; 5504pp; English.  
XX  
XX The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein  
CC described above; (15) methods of diagnosing the presence of a tumour in a  
CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.  
XX  
SQ Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;

Query Match  
Best Local Similarity 100.0%; Score 327; DB 12; Length 506;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 ATGGCGGAGCGATGGATGTGATACCCGAGCGGACCAACAGCGCGCGGCAAGAG 60  
6 ATGGCGGAGCGATGGATGTGATACCCGAGCGGACCAACAGCGCGCGGCAAGAG 65  
61 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120  
66 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 125  
121 AACTGTGCCATCTGCAGGAGCAACACATATGATCTTTGCATAGATGTCAAGCTAACCCAG 180  
126 AACTGTGCCATCTGCAGGAGCAACACATATGATCTTTGCATAGATGTCAAGCTAACCCAG 185  
181 GCGTCGCGTACTTTCAGAGAGGTGATGTCGATCGGAGGTCTGTAACCATGCTTTTTCAC 240

Db 186 GCGTCGCGTACTTTCAGAGAGGTGATGTCGATGGGAGTCTGTAACCATGCTTTTTCAC 245  
Qy 241 TTCCACTGTCATCTCTCGCTGGCTCAAAACAGCAGCAGGTGTGTCATTGGACACAGAGAG 300  
Db 246 TTCCACTGTCATCTCTCGCTGGCTCAAAACAGCAGCAGGTGTGTCATTGGACACAGAGAG 305  
Qy 301 TGGGAATTCCTCAAAAGATATGGGCACTAG 327  
Db 306 TGGGAATTCCTCAAAAGATATGGGCACTAG 332  
RESULT 4  
ADQ87156  
ID ADQ87156 standard; cDNA; 506 BP.  
XX  
AC ADQ87156;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE Human tumour-associated antigenic target (TAT) cDNA sequence #4032.  
XX  
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
KW cancer; cell proliferative disorder; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2004060270-A2.  
XX  
PD 22-JUL-2004.  
XX  
XX 15-OCT-2003; 2003WO-US029126.  
XX  
PR 18-OCT-2002; 2002US-0418988P.  
XX  
PA (GETH ) GENENTECH INC.  
PA (WUTD/) WU T D.  
PA (ZHOU/) ZHOU Y.  
XX  
XX Wu TD, Zhou Y;  
XX  
XX WPI; 2004-534300/51.  
XX  
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.  
XX  
XX Claim 1; SEQ ID NO 4032; 5504pp; English.  
XX  
XX The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the

CC presence of a protein in a sample suspected of containing the protein  
 CC described above; (15) methods of diagnosing the presence of a tumour in a  
 CC mammal; (16) a method for treating or preventing a cell proliferative  
 CC disorder associated with increased expression or activity of the above  
 CC protein; and (17) a method of binding an antibody, oligopeptide or  
 CC organic molecule to a cell that expresses the protein described above.  
 CC The TAT sequences have cytostatic activities, and can be used in gene  
 CC therapy. The composition and methods are useful for diagnosing,  
 CC preventing or treating cancer. The composition is also used for preparing  
 CC a medicament for the therapeutic treatment or diagnostic detection of a  
 CC cell proliferative disorder or cancer. The present sequence represents a  
 CC human TAT cDNA sequence from the present invention.

XX Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 327; DB 12; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-101;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCGCAAGAAG 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCGCAAGAAG 65  
 QY 61 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 66 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 125  
 QY 121 AACTGTGCATCTGCAGAACCAACATATGGATCTTTGGCATAGATGTCAAGCTAACACAG 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 126 AACTGTGCATCTGCAGAACCAACATATGGATCTTTGGCATAGATGTCAAGCTAACACAG 185  
 QY 181 GCGTCGCGTACTTTCAGAGAGTGTCTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 186 GCGTCGCGTACTTTCAGAGAGTGTCTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 245  
 QY 241 TTCCACTGATCTCTCGCTGGCTTCAAAACACAGAGTGTGTCCATTGGACAAACAGAGAG 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 246 TTCCACTGATCTCTCGCTGGCTTCAAAACACAGAGTGTGTCCATTGGACAAACAGAGAG 305  
 QY 301 TGGGAATTCAAAAGTATGGCACTAG 327  
 DB 306 TGGGAATTCAAAAGTATGGCACTAG 332

RESULT 5  
 ADQ84881  
 ID ADQ84881 standard; cDNA; 506 BP.  
 AC ADQ84881;  
 DT 07-OCT-2004 (first entry)  
 XX Human tumour-associated antigenic target (TAT) cDNA sequence #1695.  
 DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
 XX human; cell proliferative disorder; gene; ss.  
 KW Homo sapiens.  
 XX OS  
 XX WO2004060270-A2.  
 PN 22-JUL-2004.  
 PD 15-OCT-2003; 2003WO-US029126.  
 XX 18-OCT-2002; 2002US-0418988P.  
 XX (GETH ) GENENTECH INC.  
 PA (WUTD/) WU T D.  
 PA (ZHOU/) ZHOU Y.  
 XX WU TD, Zhou Y;  
 PI WU TD, Zhou Y;  
 XX

DR WPI; 2004-534300/51.  
 XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
 PT preventing or treating cell proliferative disorders such as cancer.  
 XX Claim 1; SEQ ID NO 1695; 5504pp; English.  
 PS The present invention describes an isolated tumour-associated antigenic  
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
 CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-  
 CC (c). Also described: (1) an expression vector comprising the above  
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
 CC a process for producing a polypeptide; (4) an isolated polypeptide  
 CC comprising: (a) an amino acid sequence encoded by any of the above  
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
 CC length coding region of the above nucleotide sequences; or (c) a sequence  
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
 CC an isolated antibody that binds to the above polypeptide; (7) a process  
 CC for producing the antibody; (8) an isolated oligopeptide that binds to  
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
 CC binding organic molecule that binds to the above polypeptide; (10) a  
 CC composition of matter comprising the above (chimeric) polypeptide,  
 CC antibody, oligopeptide or TAT binding organic molecule, in combination  
 CC with a carrier; (11) an article of manufacture comprising a container and  
 CC the composition of matter contained within the container; (12) methods of  
 CC inhibiting the growth of a cell that expresses the above protein, where  
 CC the growth of the cell is at least in part dependent upon a growth  
 CC potentiating effect of the above protein; (13) a method of  
 CC therapeutically treating a mammal having a cancerous tumour comprising  
 CC cells that express the above protein; (14) a method of determining the  
 CC presence of a protein in a sample suspected of containing the protein in a  
 CC mammal; (15) methods of diagnosing the presence of a tumour in a  
 CC disorder associated with increased expression or activity of the above  
 CC protein; and (17) a method of binding an antibody, oligopeptide or  
 CC organic molecule to a cell that expresses the protein described above.  
 CC The TAT sequences have cytostatic activities, and can be used in gene  
 CC therapy. The composition and methods are useful for diagnosing,  
 CC preventing or treating cancer. The composition is also used for preparing  
 CC a medicament for the therapeutic treatment or diagnostic detection of a  
 CC cell proliferative disorder or cancer. The present sequence represents a  
 CC human TAT cDNA sequence from the present invention.

XX Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;  
 SQ Query Match 100.0%; Score 327; DB 13; Length 506;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-101;  
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCGCAAGAAG 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 6 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCGCAAGAAG 65  
 QY 61 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 66 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 125  
 QY 121 AACTGTGCATCTGCAGAACCAACATATGGATCTTTGGCATAGATGTCAAGCTAACACAG 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 126 AACTGTGCATCTGCAGAACCAACATATGGATCTTTGGCATAGATGTCAAGCTAACACAG 185  
 QY 181 GCGTCGCGTACTTTCAGAGAGTGTCTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 240  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 186 GCGTCGCGTACTTTCAGAGAGTGTCTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 245  
 QY 241 TTCCACTGATCTCTCGCTGGCTTCAAAACACAGAGTGTGTCCATTGGACAAACAGAGAG 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 246 TTCCACTGATCTCTCGCTGGCTTCAAAACACAGAGTGTGTCCATTGGACAAACAGAGAG 305  
 QY 301 TGGGAATTCAAAAGTATGGCACTAG 327



Db		 306 TGGGAATTCCAAAGTAGTGGGCACTAG 332 
RESULT 6		
AAA74978		
ID	AAA74978 standard; DNA; 508 BP.	
XX AC	AAA74978;	
XX DT	02-JAN-2001 (first entry)	
DE DE	DNA encoding a human cullin-interacting RING-H2 finger protein (Rbx1).	
XX KW	Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;	
XX KW	tumour suppressor; carcinoma; Ring box associated carcinoma;	
KW KW	von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;	
KW KW	cerebellar hemangioblastoma; hemangioma; retinal angiomas;	
XX KW	pheochromocytomas; ss.	
OS OS	Homo sapiens.	
XX FH	Key Location/Qualifiers	
FT CDS	7..333	
FT FT	/*tag= a	
FT FT	/product= "cullin-interacting RING-H2 finger protein	
XX FT	(Rbx1)"	
XX PN	WO2000050445-A1.	
XX PD	31-AUG-2000.	
XX PP	25-FEB-2000; 2000WO-US004838.	
XX PR	26-FEB-1999; 99US-0121787P.	
XX PA	(OKLA-) OKLAHOMA MEDICAL RES FOUND.	
XX PI	Conaway JA, Conaway RC, Kamura T;	
DR DR	WPI; 2000-572067/53.	
DR DR	P-PSDB; AAB08813.	
XX PT	Cullin interacting RING-H2 finger protein, a component of von Hippel-	
PT PT	Lindau tumor suppressor complex and Skp1-Cdc53p-P-box protein (SCF)	
PT PT	ubiquitin ligase, useful for diagnosing and treating Ring box protein	
XX PT	associated carcinomas.	
XX PS	Claim 3; Page 35; 37pp; English.	
XX CC	The present sequence encodes a human cullin-interacting RING-H2 finger	
CC CC	protein (Ring box protein), designated Rbx1. The polypeptide is a tumour	
CC CC	suppressor. Rbx1 is useful for diagnosing a predisposition of a patient	
CC CC	to certain carcinomas. It is also useful for treating Ring box protein	
CC CC	associated carcinomas or augmenting metabolically deficient system in	
CC CC	animals. Rbx1 is also useful for evaluating the effectiveness of a	
CC CC	therapeutic treatment for Ring box associated carcinomas. Rbx1 can be	
CC CC	used to screen for agents which augment or inhibit the activity of other	
CC CC	cullin-containing ubiquitin ligase and of the VHL (von Hippel-Lindau)	
CC CC	complex controlling the conjugation of ubiquitin or ubiquitin-like	
CC CC	proteins to various sets of target proteins. Carcinomas which may be	
CC CC	treated include renal carcinomas, cerebellar hemangioblastomas and	
XX CC	hemangiomas, retinal angiomas and pheochromocytomas	
SQ SQ	Sequence 508 BP; 126 A; 106 C; 124 G; 152 T; 0 U; 0 Other;	
Query Match	100.0%; Score 327; DB 3; Length 508;	
Best Local Similarity	100.0%; Pred. No. 4.4e-101;	
Matches 327; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGCCGCAGCGATGTGGATTACCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60	
Db	7 ATGCCGCAGCGATGTGGATTACCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 66	





ID ADQ92179 standard; DNA; 503 BP.  
 XX ADQ92179;  
 AC  
 XX 07-OCT-2004 (first entry)  
 DT  
 XX Human autoantigen DNA fragment MPMGP800L05536.  
 DE  
 XX ds; autoantigen; antibody; hybridoma; biosensor chip;  
 KW extracorporeal differential diagnosis; autoimmune disease;  
 KW ribosomal protein; tubulin;  
 KW dolichyl-diphospho-oligosaccharide-glycosyl transferase;  
 KW multiple sclerosis; rheumatoid arthritis; epitope mapping;  
 KW affinity chromatography; electrophoresis; autoantibody apheresis;  
 KW RNA interference; RNAi.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004058972-A1.  
 PN  
 XX 15-JUL-2004.  
 PD  
 XX 23-DEC-2002; 2002WO-EP014731.  
 PF  
 XX 23-DEC-2002; 2002WO-EP014731.  
 PR  
 XX (THIE/) THIESEN H.  
 PA (LORE/) LORENZ P.  
 XX  
 XX Thiesen H, Lorenz P;  
 PI  
 XX WPI; 2004-543459/52.  
 DR  
 XX New human DNA autoantigens, useful as assay, diagnostic, and prognostic  
 PT reagents and for treating autoimmune disease, also related expression  
 PT products and antibodies with similar uses.  
 PT  
 XX Claim 1; SEQ ID NO 160; 110pp; German.  
 PS  
 XX This invention describes novel human DNA autoantigens which are used to  
 CC produce recombinant expression vectors; prokaryotic or eukaryotic cells;  
 CC poly- or mono-clonal antibodies (Ab) specific; hybridomas that express  
 CC monoclonal Ab; biosensor chips having an addressable sequence pattern as  
 CC probes; medical or diagnostic instruments that include the biosensor; for  
 CC extracorporeal differential diagnosis of autoimmune diseases and  
 CC predisposition to them. The autoantigen polynucleotides encode ribosomal  
 CC proteins; tubulins; dolichyl-diphospho-oligosaccharide-glycosyl  
 CC transferases and proteins. The antibodies may be labelled conventionally  
 CC with radioisotopes, coloured or fluorescent groups, or a member of the  
 CC biotin/avidin pair, or colloidal gold. The autoantigens can be directed  
 CC against mitochondria, liver-kidney microsomes; histidyl-tRNA; nuclear  
 CC membrane; neutrophil/cytoplasm; insect cells; epidermal intracellular  
 CC or basal membrane antigens; Golgi or cell nuclei, or associated with  
 CC multiple sclerosis or rheumatoid arthritis. They are useful for epitope  
 CC mapping; in affinity chromatography or electrophoresis; for diagnosis,  
 CC prognosis, control of treatment or therapeutic response of autoimmune  
 CC diseases, particularly in vitro differential diagnosis of autoimmune  
 CC diseases; to produce biosensor chips or for autoantibody apheresis.  
 CC Autoantigen DNA can be used for therapeutic RNA interference against  
 CC autoantibodies. Biochips that carry the new materials are useful in  
 CC medical and diagnostic instruments. ADQ92020-ADQ92280 represent human  
 CC autoantigens.  
 XX  
 SQ Sequence 503 BP; 131 A; 102 C; 120 G; 150 T; 0 U; 0 Other;  
 Query Match 97.9%; Score 320; DB 12; Length 503;  
 Best Local Similarity 100.0%; Pred. NO. 1.1e-98;  
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 8 CAGCGATGGATGGTATACCCGAGCGGCACCAACAGCGCGCGGCAAGCGCTTTG 67  
 1 CAGCGATGGATGGTATACCCGAGCGGCACCAACAGCGCGCGGCAAGCGCTTTG 60

68 AAGTGAAGAAAGTGAATGAGTACGAGTACCCCTCTGGCCCTGGGATATTGTTGTAATACTGTG 127  
 61 AAGTGAAGAAAGTGAATGAGTACGAGTACCCCTCTGGCCCTGGGATATTGTTGTAATACTGTG 120  
 128 CCATCTGCAGGAACACACATTTATGGATCTTTGATAGAATGTCAAGCTAACCAAGGCGTCCG 187  
 121 CCATCTGCAGGAACACACATTTATGGATCTTTGATAGAATGTCAAGCTAACCAAGGCGTCCG 180  
 188 CTACTTTCAAGAGAGTGTACTGTGCGATGGGAGTCTGTAACCATCTTTTCACTTCCACT 247  
 181 CTACTTTCAAGAGAGTGTACTGTGCGATGGGAGTCTGTAACCATCTTTTCACTTCCACT 240  
 248 GCATCTCTCGTGGCTCAAAACACACAGAGTGTGTCATTTGGACAAACAGAGAGTGGGAT 307  
 241 GCATCTCTCGTGGCTCAAAACACACAGAGTGTGTCATTTGGACAAACAGAGAGTGGGAT 300  
 308 TCCAAAAGTATGGCACTAG 327  
 301 TCCAAAAGTATGGCACTAG 320

RESULT 12  
 ABV25615/c  
 ID ABV25615 standard; cDNA; 4476 BP.  
 XX  
 AC ABV25615;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 25606.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 XX 20-FEB-2001; 2001WO-US005171.  
 PF  
 XX 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 5119-5120; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

```
XX
SQ Sequence 4476 BP; 1085 A; 1129 C; 1015 G; 1247 T; 0 U; 0 Other;
Query Match 94.4%; Score 308.8; DB 5; Length 4476;
Best Local Similarity 99.4%; Pred. No. 2.2e-94;
Matches 310; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 GCGCGCAGGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCAAGAGCG 62
DB 1018 GCCCGCAGGATGGATGGATACCCCGAGCGGACCAACAGCGCGCGGCAAGAGCG 959
QY 63 CTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 122
DB 958 CTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 899
QY 123 CTGTGCCATCTGCAGGAACCAATTATGATCTTTGTCATAGATCTCAAGCTAACCGGC 182
DB 898 CTGTGCCATCTGCAGGAACCAATTATGATCTTTGTCATAGATCTCAAGCTAACCGGC 839
QY 183 GTCCGCTACTTCAGAAAGTGTACTGTGCATGCGGAGTCTGTAAACCATGCTTTTCACTT 242
DB 838 GTCCGCTACTTCAGAAAGTGTACTGTGCATGCGGAGTCTGTAAACCATGCTTTTCACTT 779
QY 243 CCACTGCATCTCTCGCTCGCTCAAAACACAGCAGGTGTCTCATTTGGACACAGAGGTG 302
DB 778 CCACTGCATCTCTCGCTCGCTCAAAACACAGCAGGTGTCTCATTTGGACACAGAGGTG 719
QY 303 GGAATTCACAAA 314
DB 718 GGAATTCACAAA 707
RESULT 13
ABX39512
ID ABX39512 standard; cDNA; 380 BP.
XX
AC ABX39512;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #4677.
XX
KW Bovine; s8; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
FI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
DR
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 4677; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
```

```
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 380 BP; 98 A; 87 C; 103 G; 92 T; 0 U; 0 Other;
Query Match 92.5%; Score 302.4; DB 8; Length 380;
Best Local Similarity 96.6%; Pred. No. 1e-92;
Matches 309; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 8 CAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCGCTTTG 67
DB 2 CAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCGCTTTG 61
QY 68 AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 127
DB 62 AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 121
QY 128 CCATCTGCAGGAACCAACATTTATGATCTTTTCATAGATGTCACAGTCAACGAGGCGTCCG 187
DB 122 CCATCTGCAGGAACCAACATTTATGATCTTTTCATAGATGTCACAGTCAACGAGGCGTCCG 181
QY 188 CTACTTCTCAGAGAGTGTACTGTCTGCATGGGAGTCTGTAAACCATGCTTTTCACTTCCACT 247
DB 182 CTACTTCTCAGAGAGTGTACTGTCTGCATGGGAGTCTGTAAACCATGCTTTTCACTTCCACT 241
QY 248 GCATCTCTCGCTCGCTCAAAACACAGCAGGTGTGTCCATTGGACACAGAGAGTGGGAAT 307
DB 242 GCATCTCTCGCTCGCTCAAAACACAGCAGGTGTGTCCATTGGACACAGAGAGTGGGAAT 301
QY 308 TCCAAAGATGGGCACCTAG 327
DB 302 TCCAAAGATGGGCACCTAG 321
RESULT 14
AAS94844/C
ID AAS94844 standard; DNA; 5347 BP.
XX
AC AAS94844;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #99 expressed during foam cell differentiation.
XX
KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200177389-A2.
```

XX	18-OCT-2001.	DE	Human Testican-1 nucleotide sequence SEQ ID NO:9.	XX
PD		XX		
XX	04-APR-2001; 2001WO-US011128.	KW	diabetic; pre-diabetic; Type 2 diabetes; antidiabetic; gene therapy;	
PF		KW	diabetes; insulin resistance; metabolic disease; human; gene; ss.	
XX	05-APR-2000; 2000US-0195106P.	XX		
XX	(INCY-) INCYTE GENOMICS INC.	OS	Homo sapiens.	
PA		XX		
XX	Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GU, Mikita T;	PN	WO2003102163-A2.	
PI	Tai J;	XX		
XX		PD	11-DEC-2003.	
XX		XX		
DR	WPI; 2002-010925/01.	PF	04-JUN-2003; 2003WO-US017825.	
XX		XX		
XX	Composition useful for diagnosis of conditions, disorders or diseases	PR	04-JUN-2002; 2002US-0385857P.	
PT	associated with atherosclerosis, comprises several polynucleotides that	PR	04-JUN-2002; 2002US-0386013P.	
PT	are differentially expressed in foam cell development.	PR	04-JUN-2002; 2002US-0386074P.	
XX		PR	04-JUN-2002; 2002US-0386107P.	
XX		PR	05-JUN-2002; 2002US-0386314P.	
XX		PR	05-JUN-2002; 2002US-0386326P.	
PS	Claim 1; Page 151-152; 315pp; English.	PR	05-JUN-2002; 2002US-0386332P.	
XX		PR	05-JUN-2002; 2002US-0386481P.	
CC	The present invention relates to the isolation of human polynucleotide	PR	05-JUN-2002; 2002US-0386512P.	
CC	sequences that are differentially expressed during foam cell	PR	05-JUN-2002; 2002US-0386513P.	
CC	differentiation. The polynucleotide sequences of the invention or a	PR	05-JUN-2002; 2002US-0386558P.	
CC	composition comprising these polynucleotides are useful as a high	PR	05-JUN-2002; 2002US-0386600P.	
CC	throughput method for detecting altered expression of one or more	PR	05-JUN-2002; 2002US-0386615P.	
CC	polynucleotides in a sample. The polynucleotides can be used in the	PR	05-JUN-2002; 2002US-0386654P.	
CC	diagnosis of disorders associated with foam cell development such as	PR	06-JUN-2002; 2002US-0386838P.	
CC	atherosclerosis, cerebral stroke, and cardiovascular disorders such as	PR	06-JUN-2002; 2002US-0386861P.	
CC	coronary artery disease. The polynucleotide sequences can also be used as	PR	06-JUN-2002; 2002US-0386944P.	
CC	PCR primers and probes. The polynucleotides of the invention are also	PR	06-JUN-2002; 2002US-0386955P.	
CC	useful in gene therapy. AAS94746-AAS95021 represent the human	PR	06-JUN-2002; 2002US-0387017P.	
CC	polynucleotide sequences of the invention which are differentially	PR	06-JUN-2002; 2002US-0387026P.	
CC	expressed during foam cell differentiation	PR	06-JUN-2002; 2002US-0387039P.	
XX		PR	20-JUN-2002; 2002US-0386865P.	
SQ	Sequence 5347 BP; 1489 A; 1277 C; 1094 G; 1487 T; 0 U; 0 Other;	XX	(META-) METABOLEX INC.	
	Query Match 92.1%; Score 301.2; DB 6; Length 5347;	PA		
	Best Local Similarity 98.7%; Pred. No. 9.4e-92;	XX	Allan B, Gregoire F, Lavan B, Moodie S, Waters S, Wong C;	
	Matches 314; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	PI		
		XX		
QY	3 GGCGGCGGCGATGGATGTGGATACCCGAGCGGCACCAACAGCGGCGC-GGGCAAGAAGC 61	DR	WPI; 2004-053469/05.	
Db	519 GCCCGGCGGCGATGGATGTGGATACCCGAGCGGCACCAACAGCGGCGGCGCAAGAAGC 460	XX	P-PSDB; ADF42704.	
QY	62 GCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGGTATA 121	PT	Identifying an agent for treating diabetic or pre-diabetic individuals	
Db	459 GCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGGTATA 400	PT	comprises contacting an agent with a polypeptide, e.g., human ceramidase,	
QY	122 ACTGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGATGTCAAGCTAACCCAGG 181	PT	and selecting an agent that modulates the expression or activity of the	
Db	399 ACTGTGCCATCTGCAGGACCAATTATGGATCTTTGCATAGATGTCAAGCTAACCCAGG 340	PT	polypeptide.	
QY	182 CGTCCGCTACTTCAGAAGAGTGTCTGTCCATGGGAGTCTGTAAACCATGCTTTTCAC 241	XX	Disclosure; SEQ ID NO 9; 209pp; English.	
Db	339 CGTCCGCTACTTCAGAAGAGTGTCTGTCCATGGGAGTCTGTAAACCATGCTTTTCAC 280	PS		
QY	242 TCACCTGCATCTCTCGTGCTCTCAAAACAGACAGAGTGTCTTCATTCGACACAGAGAGT 301	XX	The present invention describes a method for identifying an agent for	
Db	279 TCACCTGCATCTCTCGTGCTCTCAAAACAGACAGAGTGTCTTCATTCGACACAGAGAGT 220	XX	treating a diabetic or pre-diabetic individual. The method comprises	
QY	302 GGGAAATCCAAAAGTATG 319	XX	contacting an agent to a mixture comprising a polypeptide encoded by a	
Db	219 GGGAAATCCAAAAGTAGG 202	XX	nucleic acid that hybridises under stringent conditions to a nucleic acid	
	RESULT 15	CC	encoding any of the 23 fully defined amino acid sequences given in the	
ID	ADF42703/C	CC	specification, and selecting an agent that modulates the expression or	
XX	ID ADF42703 standard; cDNA; 3484 BP.	CC	activity of the polypeptide. Also described: (1) a method of treating a	
XX	ADF42703;	CC	diabetic or pre-diabetic animal, comprising administering to the animal a	
XX		CC	therapeutic amount of an agent identified by the method described above;	
DT	26-FEB-2004 (first entry)	CC	(2) a method of introducing an expression cassette into a cell,	
XX		CC	comprising introducing into the cell an expression cassette comprising a	
		CC	promoter operably linked to a polynucleotide encoding a polypeptide,	
		CC	where the polynucleotide hybridises under stringent conditions to a	
		CC	nucleic acid encoding the above amino acid sequences; and (3) a method of	
		CC	diagnosing an individual who has Type 2 diabetes or is pre-diabetic,	
		CC	comprising detecting in a sample from the individual the level of a	
		CC	polypeptide or the level of the above-mentioned polynucleotide encoding	
		CC	the polypeptide, where a modulated level of the polypeptide or	
		CC	polynucleotide in the sample compared to a level of the polypeptide or	
		CC	polynucleotide in either a lean individual or a previous sample from the	
		CC	individual indicates that the individual is diabetic or pre-diabetic. The	
		CC	method is useful in diagnosing and treating diabetes, insulin resistance	
		CC	or related metabolic diseases in human subjects. The method may also be	

CC used in identifying agents for treating diabetic or pre-diabetic  
CC individuals. The present sequence is used in the exemplification of the  
CC present invention.

XX  
SQ Sequence 3484 BP; 894 A; 910 C; 808 G; 872 T; 0 U; 0 Other;

Query Match 90.8%; Score 296.8; DB 12; Length 3484;  
Best Local Similarity 99.3%; Pred. No. 2.4e-90;  
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3	GGCGGCGAGTGGATGGGATACCCGAGCGGCGCACCAACAGCGGCGGGCAAGAAGCG	62
Db	300	GGCGGCGAGTGGATGGGATACCCGAGCGGCGCACCAACAGCGGCGGGCAAGAAGCG	241
Qy	63	CTTTGAAGTGAAAAAGTGGATGCGAGTAGCCCTCTGGGCTGGGATATTGTGGTTGATAA	122
Db	240	CTTTGAAGTGAAAAAGTGGATGCGAGTAGCCCTCTGGGCTGGGATATTGTGGTTGATAA	181
Qy	123	CTGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGAGTGTCAAGCTAACCCAGGC	182
Db	180	CTGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGAGTGTCAAGCTAACCCAGGC	121
Qy	183	GTCCGCTACTTCAGAAGAGTGTACTGTGCGATGGGAGTCTGTAACCATGCTTTTCACTT	242
Db	120	GTCCGCTACTTCAGAAGAGTGTACTGTGCGATGGGAGTCTGTAACCATGCTTTTCACTT	61
Qy	243	CCACTGCACTCTCTCGTGGCTCAAAACACAGCAGGTGTGTCATTGGACAAACAGAGAGTG	302
Db	60	CCACTGCACTCTCTCGTGGCTCAAAACACAGCAGGTGTGTCATTGGACAAACAGAGAGTG	1

Search completed: June 24, 2005, 05:21:51  
Job time : 365 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2005, 05:57:28 ; Search time 1566 Seconds  
(without alignments)  
1305.185 Million cell updates/sec

Title: US-09-541-462B-1

Perfect score: 327

Sequence: 1 atggggcagcagtgatgt.....tccaaaagtatggcactag 327

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10J\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	327	100.0	433	17	US-10-242-535A-43377
2	327	100.0	433	18	US-10-085-783A-43377
3	327	100.0	453	17	US-10-242-535A-35025
4	327	100.0	453	18	US-10-085-783A-35025
5	327	100.0	467	17	US-10-242-535A-39933
6	327	100.0	467	18	US-10-085-783A-39933
7	327	100.0	471	17	US-10-242-535A-57254

8	327	100.0	471	18	US-10-085-783A-57254	Sequence 57254, A
9	327	100.0	472	17	US-10-242-535A-56068	Sequence 56068, A
10	327	100.0	472	18	US-10-085-783A-56068	Sequence 56068, A
11	327	100.0	508	21	US-10-913-937-3	Sequence 3, Appli
12	327	100.0	523	17	US-10-242-535A-46292	Sequence 46292, A
13	327	100.0	523	18	US-10-085-783A-46292	Sequence 46292, A
14	325	99.4	476	10	US-09-918-995-17191	Sequence 17191, A
15	321.8	98.4	4543	14	US-10-198-846-11311	Sequence 11311, A
16	316	96.6	430	17	US-10-242-535A-54751	Sequence 54751, A
17	316	96.6	430	18	US-10-085-783A-54751	Sequence 54751, A
18	308.8	94.4	4476	20	US-10-357-930-25604	Sequence 25604, A
19	302.4	92.5	380	9	US-09-960-352-4577	Sequence 4677, Ap
20	301.2	92.1	5347	16	US-10-240-365-99	Sequence 99, Appl
21	296.8	90.8	3484	20	US-10-723-860-1383	Sequence 1383, Ap
22	296.8	90.8	5111	15	US-10-205-823-382	Sequence 382, App
23	296.8	90.8	5371	20	US-10-723-860-5852	Sequence 5852, Ap
24	290.2	88.7	504	21	US-10-913-937-5	Sequence 5, Appli
25	281	85.9	468	17	US-10-242-535A-47556	Sequence 47556, A
26	281	85.9	468	18	US-10-085-783A-47556	Sequence 47556, A
27	278.6	85.2	486	17	US-10-242-535A-52747	Sequence 52747, A
28	278.6	85.2	486	18	US-10-085-783A-52747	Sequence 52747, A
29	273.8	83.7	475	17	US-10-242-535A-50604	Sequence 50604, A
30	273.8	83.7	475	18	US-10-085-783A-50604	Sequence 50604, A
31	253.6	77.6	3208	9	US-09-780-016-27	Sequence 27, Appl
32	253.6	77.6	3208	14	US-10-214-811-27	Sequence 27, Appl
33	253.6	77.6	3208	19	US-10-766-074-27	Sequence 27, Appl
34	249	76.1	300	17	US-10-242-535A-48516	Sequence 48516, A
35	249	76.1	300	18	US-10-085-783A-48516	Sequence 48516, A
36	239	73.1	464	17	US-10-242-535A-58211	Sequence 58211, A
37	239	73.1	464	18	US-10-085-783A-58211	Sequence 58211, A
38	213.4	65.3	370	17	US-10-242-535A-19847	Sequence 19847, A
39	213.4	65.3	370	18	US-10-085-783A-19847	Sequence 19847, A
40	208.8	63.9	273	17	US-10-242-535A-46883	Sequence 46883, A
41	208.8	63.9	273	18	US-10-085-783A-46883	Sequence 46883, A
42	205.8	62.9	439	10	US-09-918-995-14771	Sequence 14771, A
43	191	58.4	1050	21	US-10-487-501-2521	Sequence 2521, Ap
44	191	58.4	1050	21	US-10-487-501-4357	Sequence 4357, Ap
45	190	58.1	608	19	US-10-767-795-687	Sequence 687, App

ALIGNMENTS

RESULT 1

US-10-242-535A-43377  
; Sequence 43377, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242.535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 43377  
; LENGTH: 433  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-43377

Query Match 100.0%; Score 327; DB 17; Length 433;  
Best Local Similarity 100.0%; Pred. No. 4,le-104;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 35025
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-35025

Query Match      100.0%; Score 327; DB 18; Length 453;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 24 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 83
QY 61 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 120
DB 84 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 143
QY 121 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 180
DB 144 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 203
QY 181 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 240
DB 204 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 263
QY 241 TTCACACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTCTCCATTGGACAACAGAGAG 300
DB 264 TTCACACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTCTCCATTGGACAACAGAGAG 323
QY 301 TGGGAATTCCTCAAAAGTATGGGCACTAG 327
DB 324 TGGGAATTCCTCAAAAGTATGGGCACTAG 350

RESULT 5
US-10-242-535A-39933
; Sequence 39933, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39933
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-39933

Query Match      100.0%; Score 327; DB 17; Length 467;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
```

```
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 20 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 61 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 120
DB 80 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 139
QY 121 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 180
DB 140 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 199
QY 181 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 240
DB 200 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 259
QY 241 TTCACACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTCTCCATTGGACAACAGAGAG 300
DB 260 TTCACACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTCTCCATTGGACAACAGAGAG 319
QY 301 TGGGAATTCCTCAAAAGTATGGGCACTAG 327
DB 320 TGGGAATTCCTCAAAAGTATGGGCACTAG 346

RESULT 6
US-10-085-783A-39933
; Sequence 39933, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39933
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-39933

Query Match      100.0%; Score 327; DB 18; Length 467;
Best Local Similarity 100.0%; Pred. No. 4.2e-104;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60
DB 20 ATGGCGGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 61 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 120
DB 80 CGCTTTGAAGTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT 139
QY 121 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 180
DB 140 AACTGTGCCATCTCGAGGAACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 199
QY 181 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 240
DB 200 GCGTCCGCTACTTTCAGAGAGTGACTGTGCGATGGGAGTCTGTAACCATGCTTTTTCAC 259
QY 241 TTCACACTGCATCTCTCGCTGGCTCAAAACACAGCAGGTGTCTCCATTGGACAACAGAGAG 300
```



OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (455)..(455)  
 OTHER INFORMATION: n is a, c, g, or t  
 US-10-242-535A-56068

Query Match 100.0%; Score 327; DB 17; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-104; Indels 0; Gaps 0;  
 Matches 327; Conservative 0; Mismatches 0;

QY 1 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60  
 DB 20 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 79  
 QY 61 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 120  
 DB 80 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 139  
 QY 121 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 180  
 DB 140 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 199  
 QY 181 GGTCCGCTACTTCAGAGAGTGTCTGCGATGGGAGTCTGTAACCATGCTTTTCAC 240  
 DB 200 GGTCCGCTACTTCAGAGAGTGTCTGCGATGGGAGTCTGTAACCATGCTTTTCAC 259  
 QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACACAGGCTGTGTCATTGGACAAACAGAGAG 300  
 DB 260 TTCCACTGCATCTCGCTGGCTCAAAACACACAGGCTGTGTCATTGGACAAACAGAGAG 319  
 QY 301 TGGGAATTCAAAAAGTATGGGCACCTAG 327  
 DB 320 TGGGAATTCAAAAAGTATGGGCACCTAG 346

## RESULT 10

US-10-085-783A-56068  
 Sequence 56068, Application US/10085783A  
 Publication No. US20040037841A1  
 GENERAL INFORMATION:  
 APPLICANT: ChondroGene Inc.  
 APPLICANT: Liew, C.C.  
 TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
 FILE REFERENCE: 4231/2002  
 CURRENT APPLICATION NUMBER: US/10/085,783A  
 CURRENT FILING DATE: 2002-02-28  
 PRIOR APPLICATION NUMBER: US 60/305,340  
 PRIOR FILING DATE: 2001-07-13  
 PRIOR APPLICATION NUMBER: US 60/275,017  
 PRIOR FILING DATE: 2001-03-12  
 PRIOR APPLICATION NUMBER: US 60/271,955  
 PRIOR FILING DATE: 2001-02-28  
 NUMBER OF SEQ ID NOS: 58994  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 56068  
 LENGTH: 472  
 TYPE: DNA  
 ORGANISM: Human  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (437)..(437)  
 OTHER INFORMATION: n is a, c, g, or t  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (455)..(455)  
 OTHER INFORMATION: n is a, c, g, or t  
 US-10-085-783A-56068

Query Match 100.0%; Score 327; DB 18; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-104; Indels 0; Gaps 0;  
 Matches 327; Conservative 0; Mismatches 0;

QY 1 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60  
 DB 20 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 79  
 QY 61 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 120  
 DB 80 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 139  
 QY 121 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 180  
 DB 140 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 199  
 QY 181 GGTCCGCTACTTCAGAGAGTGTCTGCGATGGGAGTCTGTAACCATGCTTTTCAC 240  
 DB 200 GGTCCGCTACTTCAGAGAGTGTCTGCGATGGGAGTCTGTAACCATGCTTTTCAC 259  
 QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACACAGGCTGTGTCATTGGACAAACAGAGAG 300  
 DB 260 TTCCACTGCATCTCGCTGGCTCAAAACACACAGGCTGTGTCATTGGACAAACAGAGAG 319  
 QY 301 TGGGAATTCAAAAAGTATGGGCACCTAG 327  
 DB 320 TGGGAATTCAAAAAGTATGGGCACCTAG 346

## RESULT 11

US-10-913-937-3  
 Sequence 3, Application US/10913937  
 Publication No. US20050019813A1  
 GENERAL INFORMATION:  
 APPLICANT: Conaway, Joan A.  
 APPLICANT: Conaway, Ronald C.  
 APPLICANT: Kamura, Takumi  
 APPLICANT: Orlanoma Medical Research Foundation  
 TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor  
 TITLE OF INVENTION: Complex and SCF Ubiquitin Ligase  
 FILE REFERENCE: 021044-004600US  
 CURRENT APPLICATION NUMBER: US/10/913,937  
 CURRENT FILING DATE: 2004-08-05  
 PRIOR APPLICATION NUMBER: US/09/914,324  
 PRIOR FILING DATE: 2001-08-22  
 PRIOR APPLICATION NUMBER: US 60/121,787  
 PRIOR FILING DATE: 1999-02-26  
 PRIOR APPLICATION NUMBER: WO PCT/US00/04838  
 PRIOR FILING DATE: 2000-02-25  
 NUMBER OF SEQ ID NOS: 12  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 3  
 LENGTH: 508  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (7)..(333)  
 OTHER INFORMATION: Rbx1  
 US-10-913-937-3

Query Match 100.0%; Score 327; DB 21; Length 508;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-104; Indels 0; Gaps 0;  
 Matches 327; Conservative 0; Mismatches 0;

QY 1 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60  
 DB 7 ATGGCGGCGAGCATGGATGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 66  
 QY 61 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 120  
 DB 67 CGCTTTGAAGTGAAAAAGTGGATGCAATGAGTGGATGCTGGGCTGGGATATTGTTGAT 126  
 QY 121 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 180  
 DB 127 AACTGTGCCATCTGCAGGACCAACATTATGATGCTTGGATAGATGTCAGCTAACCCAG 186



US-09-918-995-17191

Query Match	99.4%	Score 325;	DB 10;	Length 476;
Best Local Similarity	100.0%;	Pred. No. 2.2e-103;		
Matches 325;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3	GGCGGCAGCGATGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCG	62	
Db	74	GGCGGCAGCGATGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCG	133	
Qy	63	CTTTGGAAGTCAAAAAGTGGAAATGCAGTAGCCCTCTGGGCGCTGGGATATTGTGGTTGATAA	122	
Db	134	CTTTGGAAGTCAAAAAGTGGAAATGCAGTAGCCCTCTGGGCGCTGGGATATTGTGGTTGATAA	193	
Qy	123	CTGTGCCATCTGCAGGAAACCAATTATGGATCTTTTGCATAGAAATGTCAAGCTAACCAAGGC	182	
Db	194	CTGTGCCATCTGCAGGAAACCAATTATGGATCTTTTGCATAGAAATGTCAAGCTAACCAAGGC	253	
Qy	183	GTCGCGTACTTTCAGAAAGATGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCACTT	242	
Db	254	GTCGCGTACTTTCAGAAAGATGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCACTT	313	
Qy	243	CCACTGCATCTCTCGCTGGCTCAAAAACAACGACAGGTGTGTCATTTGGACAACAGAGAGTG	302	
Db	314	CCACTGCATCTCTCGCTGGCTCAAAAACAACGACAGGTGTGTCATTTGGACAACAGAGAGTG	373	
Qy	303	GGAAATCCAAAAGTATGGCACTAG	327	
Db	374	GGAAATCCAAAAGTATGGCACTAG	398	

**RESULT 15**

```

US-10-198-846-11311/c
; Sequence 11311, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11311
; LENGTH: 4543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-11311

```

Query Match	98.4%	Score 321.8	DB 14	Length 4543
Best Local Similarity	99.4%	Pred. No. 7.5e-102		
Matches 323	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy	3	GGCGGCAGCGATCGATCTGGATACCCGAGCGGCACCAACAGCGCGCGGCGCAGAGCG	62	
Db	1089	GCCTCGAGCGATGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCGCAGAGCG	1030	
Qy	63	CTTTGAAGTGAAAAAGTGGAAATCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAA	122	
Db	1029	CTTTGAAGTGAAAAAGTGGAAATCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAA	970	
Qy	123	CTGTGCCATCTGCAGGAACACCATATTAGATCTTTGCATAGATGTCAAGCTAACCAAGCG	182	
Db	969	CTGTGCCATCTGCAGGAACACCATATTAGATCTTTGCATAGATGTCAAGCTAACCAAGCG	910	
Qy	183	GTCGGCTACTTTCAGAACAGTGTACTGTGCGCATGGGGAGTCTGTAAACCATGCTTTTCACTT	242	

[illegible]

Search completed: June 24, 2005, 08:54:11  
Job time : 1568 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2005, 00:30:07 ; Search time 2179 Seconds  
(without alignments)  
5712.261 Million cell updates/sec

Title: US-09-541-462B-1  
Perfect score: 327  
Sequence: 1 atggcggcagcagtgatgt.....tccaaaagatggcactag 327

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hlc:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gssi:  
9: gb\_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	462	7	CN305890 170005830
2	327	100.0	471	4	BG339057 602436882
3	327	100.0	473	5	BX283972 BX283972
4	327	100.0	488	4	BG777485 602664820
5	327	100.0	498	6	CB296618 12B22005
6	327	100.0	509	7	CN305892 170006001
7	327	100.0	522	1	AL711573 DKFp6861
8	327	100.0	525	4	BG337472 602435003
9	327	100.0	527	4	BM715215 UI-B-CL1
10	327	100.0	527	5	BG932751 AGENCOURT
11	327	100.0	527	6	CB999822 AGENCOURT
12	327	100.0	529	4	BM757406 K-EST0036
13	327	100.0	531	6	CD523457 AGENCOURT
14	327	100.0	533	5	BUT29963 UI-B-CK1
15	327	100.0	537	4	BG478622 602525509
16	327	100.0	545	5	BG941262 AGENCOURT
17	327	100.0	546	4	BI198015 602762132
18	327	100.0	546	5	BG601181 AGENCOURT
19	327	100.0	548	1	AV716338 AV716338
20	327	100.0	548	5	BG601301 AGENCOURT
21	327	100.0	549	4	BI958425 603386437
22	327	100.0	550	5	BG955646 AGENCOURT
23	327	100.0	554	6	CD249363 AGENCOURT
24	327	100.0	555	5	BUS32840 AGENCOURT

25	327	100.0	555	5	BUS601080	BUS601080 AGENCOURT
26	327	100.0	557	5	BUS94591	BUS94591 AGENCOURT
27	327	100.0	558	6	CD388268	CD388268 AGENCOURT
28	327	100.0	559	7	CK003869	CK003869 AGENCOURT
29	327	100.0	562	6	CD175085	CD175085 AGENCOURT
30	327	100.0	563	5	BUS61191	BUS61191 AGENCOURT
31	327	100.0	570	5	BUS97842	BUS97842 AGENCOURT
32	327	100.0	573	5	BUS33405	BUS33405 AGENCOURT
33	327	100.0	581	5	BF308845	BF308845 AGENCOURT
34	327	100.0	583	5	BF317451	BF317451 AGENCOURT
35	327	100.0	616	4	BG705958	BG705958 AGENCOURT
36	327	100.0	736	4	BG503311	BG503311 AGENCOURT
37	327	100.0	801	5	BUS33420	BUS33420 AGENCOURT
38	327	100.0	845	5	BUS95949	BUS95949 AGENCOURT
39	327	100.0	886	4	BG481544	BG481544 AGENCOURT
40	327	100.0	904	4	BM459634	BM459634 AGENCOURT
41	327	100.0	945	5	BQ216738	BQ216738 AGENCOURT
42	327	100.0	1098	5	BQ050461	BQ050461 AGENCOURT
43	326	99.7	403	4	BG531117	BG531117 AGENCOURT
44	326	99.7	448	2	BE738587	BE738587 AGENCOURT
45	325.4	99.5	405	4	BF956282	BF956282 AGENCOURT

ALIGNMENTS

RESULT 1  
CN305890 462 bp mRNA linear EST 16-MAY-2004  
LOCUS 17000583044616 GRN\_PREHEP Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN305890.1 GI:47322304  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 462)  
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.  
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation  
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
COMMENT Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 462 Std Error: 0.00.

FEATURES  
source

1.462  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, DMSO-treated H9 cell line"  
/clone\_lib="GRN\_PREHEP"  
/note="oligo dt primed, full-length enriched cDNA library from DMSO-treated HES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

Query Match 100.0%; Score 327; DB 7; Length 462;  
Best Local Similarity 100.0%; Pred. No. 3.3e-91;  
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGCGAGTGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAG 60  
|||||  
DB 23 ATGGCGGCGAGTGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCGGCAAGAG 82  
|||||

```

QY 61 CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGGTTGAT 120
DB 83 CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGGTTGAT 142
QY 121 AACTGTGCCATCTGCAGGAAACCAATATGGATCTTTGGCATAGAAATGTCAAGCTAACACG 180
DB 143 AACTGTGCCATCTGCAGGAAACCAATATGGATCTTTGGCATAGAAATGTCAAGCTAACACG 202
QY 181 GCGTCGCTACTTCAGAGAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 240
DB 203 GCGTCGCTACTTCAGAGAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 262
QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTGTCATTTGGACAACACAGAG 300
DB 263 TTCCACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTGTCATTTGGACAACACAGAG 322
QY 301 TGGGAATTCAAAAGTATGGCACTAG 327
DB 323 TGGGAATTCAAAAGTATGGCACTAG 349

RESULT 2
BG339057
LOCUS 602436882F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4554597 5',
DEFINITION mRNA sequence.
ACCESSION BG339057
VERSION BG339057.1 GI:13145495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 471)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgapps@email.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LICM1254 row: d column: 22
High quality sequence stop: 470.
FEATURES
source
1..471
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4554597"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/notes="Organ: uterus; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 100.0%; Score 327; DB 4; Length 471;
Best Local Similarity 100.0%; Pred. NO. 3.3e-91;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGCGAGCGATGGATGTGGATATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 60

```

```

DB 42 ATGGCGGCGAGCGATGGATGTGGATATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 101
QY 61 CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGGTTGAT 120
DB 102 CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGGTTGAT 161
QY 121 AACTGTGCCATCTGCAGGAAACCAATATGGATCTTTGGCATAGAAATGTCAAGCTAACACG 180
DB 162 AACTGTGCCATCTGCAGGAAACCAATATGGATCTTTGGCATAGAAATGTCAAGCTAACACG 221
QY 181 GCGTCGCTACTTCAGAGAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 240
DB 222 GCGTCGCTACTTCAGAGAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 281
QY 241 TTCCACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTGTCATTTGGACAACACAGAG 300
DB 282 TTCCACTGCATCTCGCTGGCTCAAAACACAGCAGGTGTGTCATTTGGACAACACAGAG 341
QY 301 TGGGAATTCAAAAGTATGGCACTAG 327
DB 342 TGGGAATTCAAAAGTATGGCACTAG 368

RESULT 3
BG339057
LOCUS 602436882F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:958P131435 ;
DEFINITION IMAGE:4652028, mRNA sequence.
ACCESSION BG339057
VERSION BG339057.1 GI:28848426
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IWAGp958P131435.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMW-M13u. Primer sequence: CGTTGTAAACGACGCCACT.
FEATURES
source
1..473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IWAGp958P131435 ; IMAGE:4652028"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/notes="Organ: placenta; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit

```





GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN	Query Match	100.0%;	Score 327;	DB 4;	Length 525;
	Best Local Similarity	100.0%;	Pred. No. 3.4e-91;		
	Matches 327;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGCGGCAGCGATGGATGTGGATACCCGAGCGGCACCAACAGCGCGCGGCGGCAAGAAG	60		
Db	7	ATGGCGGCAGCGATGGATGTGGATACCCGAGCGGCACCAACAGCGCGCGGCGGCAAGAAG	66		
Qy	61	CGCTTTGAAGTGAAGAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT	120		
Db	67	CGCTTTGAAGTGAAGAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGAT	126		
Qy	121	AACTGTGCCATCTGCAGGAACCAACATTATGGATCTTTGCATAGAAATGTCNAGCTAACCCAG	180		
Db	127	AACTGTGCCATCTGCAGGAACCAACATTATGGATCTTTGCATAGAAATGTCNAGCTAACCCAG	186		
Qy	181	GGCTCGCGCTACTTTCAAGACAGTGTACTGTCCGATGGGAGTCTGTAAACCATGCTTTTCAC	240		
Db	187	GGCTCGCGCTACTTTCAAGACAGTGTACTGTCCGATGGGAGTCTGTAAACCATGCTTTTCAC	246		
Qy	241	TTTCCATGTCATCTCTCGCTGGGCTCAAAACACGACAGGTGTGTCCATTGGACAAACAGAGAG	300		
Db	247	TTTCCATGTCATCTCTCGCTGGGCTCAAAACACGACAGGTGTGTCCATTGGACAAACAGAGAG	306		
Qy	301	TGGGAATTCAAAAGTATGGGCACTAG	327		
Db	307	TGGGAATTCAAAAGTATGGGCACTAG	333		

RESULT 9	BM715215	527 bp	mRNA	linear	EST 28-FEB-2002
LOCUS	BM715215				
DEFINITION	UI-E-CL1-afi-h-19-0-UI.r2	UI-E-CL1	Homo sapiens	CDNA clone	
	UI-E-CL1-afi-h-19-0-UI.5'			mRNA sequence.	
ACCESSION	BM715215				
VERSION	BM715215.1	GI:19028473			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 527)				
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.				
TITLE	Normalization and subtraction: two approaches to facilitate gene				
	discovery				
JOURNAL	Genome Res. 6 (9), 791-806 (1996)				
MEDLINE	97044477				
PUBMED	8889548				
COMMENT	Contact: Soares, MB				
	Coordinated Laboratory for Computational Genomics				
	University of Iowa				
	375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA				
	Tel: 319 335 8250				
	Fax: 319 335 9565				
	Email: bento-soares@uiowa.edu				
	Tissue Procurement: Dr. Gregg Hageman				
	CDNA Library preparation: Dr. M. Bento Soares, University of Iowa				
	CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa				
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa				
	Clone Distribution: Researchers may obtain clones from Research				
	Genetics (www.resgen.com).				
	Seq primer: M13 Reverse				
FEATURES	Location/Qualifiers				
source	1..527				
	/organism="Homo sapiens"				

```
/clone lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiA; Site_2: SfiB;
cDNA-collection"
```

Query Match	100.0%	Score 327;	DB 1;	Length 522;
Best Local Similarity	100.0%	Pred. No. 3.4e-91;		
Matches 327; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ATGCGCGCAGCGATGGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG	60	
DB	21	ATGCGCGCGGATGGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG	80	
QY	61	CGCTTTGAAGTGAAAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGTTGAT	120	
DB	81	CGCTTTGAAGTGAAAAAGTGGAAATGCAGTAGCCCTCTGGGCCTGGGATATTGTGTTGAT	140	
QY	121	AACGTGCCATCTGCAGAACCAATTATGGATCTTTGCATAGAAATGTCAAGCTTAACAG	180	
DB	141	AACGTGCCATCTGCAGAACCAATTATGGATCTTTGCATAGAAATGTCAAGCTTAACAG	200	
QY	181	CGGTCCGCTACTTCAGAAGAGTGACTGTCCCATGGGAGTCTGTAAACATGCTTTTCAC	240	
DB	201	CGGTCCGCTACTTCAGAAGAGTGACTGTCCCATGGGAGTCTGTAAACATGCTTTTCAC	260	
QY	241	TTCCACTTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACAAGAGAG	300	
DB	261	TTCCACTTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACAAGAGAG	320	
QY	301	TGGGAATTCCAAAAGTATGGGCACCTAG	327	
DB	321	TGGGAATTCCAAAAGTATGGGCACCTAG	347	

RESULT 8	525 bp	mRNA	linear	EST 27-FEB-2001
BG337472	602435003F1	NIH_MGC_46	Homo sapiens	CDNA clone IMAGE:4553064 5',
LOCUS	mRNA sequence.			
DEFINITION	BG337472			
ACCESSION	BG337472			
VERSION	BG337472.1	GI:13143910		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 525)			
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rsapbs@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LbCMi250 row: e column: 01 High quality sequence stop: 525. Location/Qualifiers 1. 525 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4553064" /tissue_type="leiomyosarcoma cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_46" /note="Organ: uterus; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5 adaptor:			
FEATURES	source			



ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS	1 (bases 1 to 527)
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits CDNA Library Preparation: CLONTECH Laboratories, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: NDCM123 row: p column: 18 High quality sequence stop: 518.

## FEATURES

```

1. 527
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30322649"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_18c"
/name="Organ: Pooled-Skin, Vector: pDNR-LIB; Site 1: Sfil
(ggcatatggcc); Site 2: Sfil (ggccgcctccggcc); library is
oligo-dt primed and directionally cloned. cDNA was
prepared from a pooled samples of tissues from Skin,
meninges, duramatter, pia matter and choroid plexus.
5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCCATATGGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGGCCGACATG-dt(30)EN-3'
(where B = A, C, G and N = A, C, G, or T). Average
insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library"

```

## ORIGIN

Query Match	100.0%	Score 327;	DB 6;	Length 527;
Best Local Similarity	100.0%;	Pred. No. 3.4e-91;		
Matches 327;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

1	ATGCGCGCAGCGATGTGGATACCCCGAGCGGCA	60
	CCAAACAGCGCGCGCGCGCGCGCGCGCGCGCG	
25	ATGCGCGCAGCGATGTGGATACCCCGAGCGGCA	84
	CCAAACAGCGCGCGCGCGCGCGCGCGCGCGCG	
61	CGCTTTGAAGTGAAGAGTGGAAATGCAATAGT	120
	AGCCCTCTGGGCCCTGGGATATTGTGGTTGAT	
85	CGCTTTGAAGTGAAGAGTGGAAATGCAATAGT	144
	AGCCCTCTGGGCCCTGGGATATTGTGGTTGAT	
121	AACGTGCGCATCTGCAGGAACCAATTATGAT	180
	CTTTTGGCATGAATGTCGAAGCTCAAGCTAAC	
145	AACGTGCGCATCTGCAGGAACCAATTATGAT	204
	CTTTTGGCATGAATGTCGAAGCTCAAGCTAAC	
181	GGTCCGCTACTTCAGAAAGTGTACTGTCCGAT	240
	TGGGGAGTCTGTGAACCATGCTTTTTCAC	
205	GGTCCGCTACTTCAGAAAGTGTACTGTCCGAT	264
	TGGGGAGTCTGTGAACCATGCTTTTTCAC	
241	TTCCACTGCATCTCTCGCTGGCTCAAAACAC	300
	AGAGTGTGTCCATTGGACCAACAGAGAG	
265	TTCCACTGCATCTCTCGCTGGCTCAAAACAC	324
	AGAGTGTGTCCATTGGACCAACAGAGAG	
301	TGGGAATTCCAAAAGTATGGGCAC	327
	CTAG	
325	TGGGAATTCCAAAAGTATGGGCAC	351
	CTAG	

RESULT 12  
BM757406

LOCUS	BM757406	529 bp	mRNA	linear	EST 04-MAR-2002
DEFINITION	K-EST0036388 S1SNUS Homo sapiens cDNA clone S1SNUS-29-H12 5', mRNA sequence.				
ACCESSION	BM757406				
VERSION	BM757406.1				
KEYWORDS	EST.				
SOURCE	GI:19087021				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1. (bases 1 to 529)				
	Kim, N.S., Hahn, Y.S., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.B., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.				
TITLE	21C Frontier Korean EST Project 2001				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Kim YS				
	Genome Research Center				
	Korea Research Institute of Bioscience & Biotechnology				
	52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea				
	Tel: +82-42-860-4470				
	Fax: +82-42-860-4409				
	Email: yongsung@mail.kribb.re.kr				
	Plate: 29 row: H column: 12				
	High quality sequence stop: 529.				

FEATURES	SOURCE
----------	--------

```

1. 529
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-29-H12"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="Top10F"
/clone_lib="S1SNUS"
/notes="Organ: Stomach; Vector:
Site 2: NotI; The poly (A) + R
bacterial alkaline phosphatase
with tabacco acid pyrophosphatase
intact mRNA was ligated with
I site by treatment of T4 RNA
ligase. The cDNA was synthesized from oli-
goprimer with 3'-tailed vector
adjusted to have about 60nt.
circularized with E. coli DNA
ligase. The cDNA was inserted into
EcoRI which site is also incl
converted to a DNA strand by
obtained cDNA vectors were used
competent cells E. coli Top10
The cDNA libraries constructed
full-length enriched cDNA lib

```

## ORIGIN

	Query Match	100.0%;	Score 327;	DB 4;	Length 529;
	Best Local Similarity	100.0%;	Pred. No. 3.4e-91;		
	Matches 327;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCGGCACGATGGATGTGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAAG	60		
Db	33	ATGGCGGCACGATGGATGTGGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAAG	92		
QY	61	CGCTTTGGAAGTGAAGTGAATGCAGTAGCCCTCTGGGCTCTGGGATATTGTGGTTGAT	120		
Db	93	CGCTTTGGAAGTGAAGTGAATGCAGTAGCCCTCTGGGCTCTGGGATATTGTGGTTGAT	152		
QY	121	AACGTGCGCATCTGCAGGAACACACATTATGGAATCTTTTGCATAGAAATGTCACGCTAACCCAG	180		
Db	153	AACGTGCGCATCTGCAGGAACACACATTATGGAATCTTTTGCATAGAAATGTCACGCTAACCCAG	212		
QY	181	GGGTCCGCTACTCTCAGAAAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTTCAC	240		

```

Db      213 GCGTCGCTACTTTCAGAAAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTCAC 272
Qy      241 TTCCACTGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCTTGGACACAGAGAG 300
Db      273 TTCCACTGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCTTGGACACAGAGAG 332
Qy      301 TGGGAATTCCAAAGTATGGCACTAG 327
Db      333 TGGGAATTCCAAAGTATGGCACTAG 359

RESULT 13
CD523457
LOCUS   CD523457
DEFINITION AGENCOURT_14360071 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:304112182 5', mRNA sequence.
ACCESSION CD523457
VERSION    CD523457.1 GI:31455175
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      NIH-MGC http://mgs.nci.nih.gov/.
JOURNAL    National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT    Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaps-x@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM205 row: c column: 07
High quality sequence stop: 518.

FEATURES             source
Location/Qualifiers
1. .531
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:304112182"
/tissue_type="pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_191"
/notes="vector: pNR-LIB; site_1: Sfil (ggccattatggcc);
site_2: Sfil (ggcgctcgcc); Library is oligo-dT primed
and directionally cloned. PBMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGCCCATATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGCGCGGGCGGCAGATG-dT(30)BN-3' (where B = A,
C, G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 100.0%; Score 327; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 3.5e-91;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGCGCGCAGGATGGATGGATACCCCGCGGCACCAACAGCGCGCGGCAAGAAG 60
Db      23 ATGCGCGCAGGATGGATGGATACCCCGCGGCACCAACAGCGCGCGGCAAGAAG 82

```

---

```

Qy      61 CGCTTTGAAGTGAAAAAGTGGAAATCAGTAGCCCTCTGGGCTCGGGAATATTGTGTTGAT 120
Db      83 CGCTTTGAAGTGAAAAAGTGGAAATCAGTAGCCCTCTGGGCTCGGGAATATTGTGTTGAT 142
Qy      121 AACTGTGCCATCTGCAGNAACCAATATTGGATCTTTCATAGAAATGTCAAGCTAACCCAG 180
Db      143 AACTGTGCCATCTGCAGNAACCAATATTGGATCTTTCATAGAAATGTCAAGCTAACCCAG 202
Qy      181 GGTCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 240
Db      203 GGTCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 262
Qy      241 TTCCACTGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCTTGGACACAGAGAG 300
Db      263 TTCCACTGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCTTGGACACAGAGAG 322
Qy      301 TGGGAATTCCAAAGTATGGCACTAG 327
Db      323 TGGGAATTCCAAAGTATGGCACTAG 349

RESULT 14
BU729963/c
LOCUS   BU729963
DEFINITION UI-E-CK1-afi-h-19-0-UI.s1 UI-E-CK1 Homo sapiens cDNA clone
IMAGE:304112182 5', mRNA sequence.
ACCESSION BU729963
VERSION    BU729963.1 GI:23653376
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 533)
JOURNAL    Donald, M.F., Lennon, G. and Soares, M.B.
COMMENT    Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES             source
Location/Qualifiers
1. .533
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-afi-h-19-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/notes="Organ: eye; Vector: pRT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 24, 2005, 06:37:52 ; Search time 3372 Seconds  
(without alignments)  
1551.946 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVTPSGTNSGAGKK.....KTRQVCPLDNRWEFQKYG 108

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09541462/runat.23062005.122624.9483/app.query.fasta\_1.263  
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09541462 @CGN 1 1 3731 @runat.23062005.122624.9483 -NCPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	327	9 AF142059	Homo sapi
2	616	100.0	433	6 CQ698451	Sequence
3	616	100.0	453	6 CQ690099	Sequence
4	616	100.0	467	6 CQ695007	Sequence

5	616	100.0	471	6 CQ712328	Sequence
6	616	100.0	472	6 CQ711142	Sequence
7	616	100.0	482	6 AX888031	Sequence
8	616	100.0	482	6 BD027641	Sequence
9	616	100.0	504	6 BD271522	VonHippel
10	616	100.0	504	10 AF140599	Mus muscu
11	616	100.0	507	6 CQ729899	Sequence
12	616	100.0	508	6 BD271520	VonHippel
13	616	100.0	508	9 AF140598	Homo sapi
14	616	100.0	514	5 BC082183	Xenopus l
15	616	100.0	523	6 CQ701366	Sequence
16	616	100.0	531	10 BC051473	Mus muscu
17	616	100.0	535	9 BC001466	Homo sapi
18	616	100.0	535	9 CR456560	Homo sapi
19	616	100.0	554	9 BC017370	Homo sapi
20	616	100.0	1612	10 BC056992	Mus muscu
21	616	100.0	1616	10 BC027396	Mus muscu
22	610	99.0	153274	2 AY555280	Mus muscu
23	605	98.2	536	5 BC081497	Danio rer
24	605	98.2	172747	2 AC112970	Mus muscu
25	604	98.1	1933	5 AY027936	Salmo sal
26	603	97.9	430	6 CQ709825	Sequence
27	603	97.9	484	5 CR390140	Gallus ga
28	603	97.9	503	6 CQ832496	Sequence
29	586	95.1	4476	6 CQ493737	Sequence
30	584	94.8	497	9 HMYQ60A05	Homo sapi
31	573	93.0	5347	6 AX281690	Sequence
32	554	89.9	3484	9 HSTEST	H. sapiens m
33	545	88.5	306	9 AY099360	Homo sapi
34	545	88.5	468	6 CQ702730	Sequence
35	535	86.9	692	6 CQ603273	Sequence
36	535	86.9	1181	3 AY119265	Drosophil
37	526.5	85.5	830	3 AK116119	Ciona int
38	526	85.4	582	8 BT009441	Triticum
39	524.5	85.1	703	8 AK119792	Oryza sat
40	524.5	85.1	721	8 AK059300	Oryza sat
41	524.5	85.1	1715	8 AK099287	Oryza sat
42	523	84.9	687	8 AY302066	Populus t
43	520.5	84.5	690	8 AK073331	Oryza sat
44	515.5	83.7	635	8 AY086913	Arabidops
45	515.5	83.7	665	8 AK118181	Arabidops

ALIGNMENTS

RESULT 1	AF142059	Homo sapiens	327 bp	mRNA	linear	PRI 14-JUL-1999
LOCUS	AF142059	Homo sapiens	327 bp	mRNA	linear	PRI 14-JUL-1999
DEFINITION	AF142059	Homo sapiens	327 bp	mRNA	linear	PRI 14-JUL-1999
ACCESSION	AF142059	Homo sapiens	327 bp	mRNA	linear	PRI 14-JUL-1999
VERSION	AF142059.1	GI:4809215				
KEYWORDS						
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						

```
/db_xref="taxon:9606"
1..327
/gene="ROC1"
CDS
1..327
/gene="ROC1"
/note="Cullin partnering protein with associated ubiquitin
ligase activity"
/codon_start=1
/product="RING finger protein"
/protein_id="AD30146.1"
/db_xref="GI:4809216"
/translation="MAAAMDVDTPSGTNSGAGKRFVKKMNAVALMAWDIYVDNCAL
CRNHIMDLICIEQANQASATSECTVAMGVNCFHFCISRWLKTRQVCLDREWE
FKYGH"

ORIGIN
Alignment Scores:
Pred. No.: 9.32e-59 Length: 327
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-541-462B-2 (1-108) x AF142059 (1-327)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
|
|
|
Db 1 ATGGCGGCGAGCGATGGATGGATATCCCGAGCGGCACCAACAGCGCGGCGCAAGAAG 60
|
|
|
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrrAlaTrrAspIleValValAsp 40
|
|
|
Db 61 CGCTTTGAAGTGAAAAGTGGATGCAATGAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT 120
|
|
|
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
|
|
|
Db 121 AACTGTGCCATCTGCAGGAACCAACATATTGATCTTTGCATAGATGTCAAGCTAACCAAG 180
|
|
|
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrrGlyValCysAsnHisAlaPheHis 80
|
|
|
Db 181 GCCTCCGCTACTTTCAGAAAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 240
|
|
|
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
|
|
|
Db 241 TTCACATGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACACACAGAG 300
|
|
|
QY 101 TrpGluPheGlnLysTyrGlyHis 108
|
|
|
Db 301 TGGGAATTCAAAAGTATGGGCAC 324
|
|
|

RESULT 2
CQ698451
LOCUS CQ698451 433 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 43377 from Patent WO02070737.
ACCESSION CQ698451
VERSION CQ698451.1 GI:4252856
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Liew,C.C., Marshall,W.E. and Zhang,H.
AUTHORS Compositions and methods relating to osteoarthritis
TITLE Patent: WO 02070737-A 43377 12-SEP-2002;
JOURNAL Chondrogene Inc. (CA)
FEATURES
source
1..433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.32e-59 Length: 433
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-541-462B-2 (1-108) x CQ698451 (1-433)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
|
|
|
Db 24 ATGGCGGCGAGCGATGGATGGATATCCCGAGCGGCACCAACAGCGCGGCGCAAGAAG 83
|
|
|
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrrAlaTrrAspIleValValAsp 40
|
|
|
```

```
Alignment Scores:
Pred. No.: 1.25e-58 Length: 433
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-541-462B-2 (1-108) x CQ698451 (1-433)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
|
|
|
Db 20 ATGGCGGCGAGCGATGGATGGATATCCCGAGCGGCACCAACAGCGCGGCGCAAGAAG 79
|
|
|
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrrAlaTrrAspIleValValAsp 40
|
|
|
Db 80 CGCTTTGAAGTGAAAAGTGGATGCAATGAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT 139
|
|
|
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
|
|
|
Db 140 AACTGTGCCATCTGCAGGAACCAACATATTGATCTTTGCATAGATGTCAAGCTAACCAAG 199
|
|
|
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrrGlyValCysAsnHisAlaPheHis 80
|
|
|
Db 200 GCCTCCGCTACTTTCAGAAAGTGTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC 259
|
|
|
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
|
|
|
Db 260 TTCACATGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACACACAGAG 319
|
|
|
QY 101 TrpGluPheGlnLysTyrGlyHis 108
|
|
|
Db 320 TGGGAATTCAAAAGTATGGGCAC 343
|
|
|

RESULT 3
CQ690099
LOCUS CQ690099 453 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 35025 from Patent WO02070737.
ACCESSION CQ690099
VERSION CQ690099.1 GI:42225827
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Liew,C.C., Marshall,W.E. and Zhang,H.
AUTHORS Compositions and methods relating to osteoarthritis
TITLE Patent: WO 02070737-A 35025 12-SEP-2002;
JOURNAL Chondrogene Inc. (CA)
FEATURES
source
1..453
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.32e-58 Length: 453
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-541-462B-2 (1-108) x CQ690099 (1-453)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
|
|
|
Db 24 ATGGCGGCGAGCGATGGATGGATATCCCGAGCGGCACCAACAGCGCGGCGCAAGAAG 83
|
|
|
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrrAlaTrrAspIleValValAsp 40
|
|
|
```

```

Db      84  CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGTTGAT 143
QY      41  AenCysAlaIleCysArgAenHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      144  AACTGTGCCATCTGCAGGAACACACATATTGGATCTTTGCATAGAAATGTCAAGCTAACCCAG 203
QY      61  AlaSerAlaThrSerGluCysThrValAlaATpGlyValCysAsnHisAlaPheHis 80
Db      204  GCGTCGCTACTTCAGAAAGTGTACTGTCGATGGGAGTCTGTAAACCATCTTTTTCAC 263
QY      81  PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      264  TTCCACTGCATCTCTCGCTGGCTCAAAACACACAGAGTGTGTCCATTGGACACACAGAGAG 323
QY      101  TrpGluPheGlnLysTyrGlyHis 108
Db      324  TGGGAATTCAAAAGTATGGGCAC 347

RESULT 4
CQ695007
LOCUS      CQ695007
DEFINITION Sequence 39933 from Patent WO02070737.
ACCESSION CQ695007
VERSION    CQ695007.1 GI:42240530
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Liew, C.C., Marshall, W.E. and Zhang, H.
TITLE      Compositions and methods relating to osteoarthritis
JOURNAL    Patent: WO 02070737-A 39933 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES   Location/Qualifiers
            source
            1..467
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      1,368-58      Length:      467
Score:          616.00      Matches:    108
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-541-462B-2 (1-108) x CQ695007 (1-467)

QY      1  MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20
Db      20  ATGGCGGCAGCATGTGATGTGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 79
QY      21  ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db      80  CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGTTGAT 139
QY      41  AenCysAlaIleCysArgAenHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      140  AACTGTGCCATCTGCAGGAACACACATATTGGATCTTTGCATAGAAATGTCAAGCTAACCCAG 199
QY      61  AlaSerAlaThrSerGluCysThrValAlaATpGlyValCysAsnHisAlaPheHis 80
Db      200  GCGTCGCTACTTCAGAAAGTGTACTGTCGATGGGAGTCTGTAAACCATCTTTTTCAC 259
QY      81  PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      260  TTCCACTGCATCTCTCGCTGGCTCAAAACACACAGAGTGTGTCCATTGGACACACAGAGAG 319
QY      101  TrpGluPheGlnLysTyrGlyHis 108
            |||||||

```

```

Db      320  TGGGAATTCAAAAGTATGGGCAC 343

RESULT 5
CQ712328
LOCUS      CQ712328
DEFINITION Sequence 57254 from Patent WO02070737.
ACCESSION CQ712328
VERSION    CQ712328.1 GI:42273185
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Liew, C.C., Marshall, W.E. and Zhang, H.
TITLE      Compositions and methods relating to osteoarthritis
JOURNAL    Patent: WO 02070737-A 57254 12-SEP-2002;
            Chondrogene Inc. (CA)
FEATURES   Location/Qualifiers
            source
            1..471
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      1,378-58      Length:      471
Score:          616.00      Matches:    108
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-541-462B-2 (1-108) x CQ712328 (1-471)

QY      1  MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20
Db      17  ATGGCGGCAGCATGTGATGTGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 76
QY      21  ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db      77  CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGTTGAT 136
QY      41  AenCysAlaIleCysArgAenHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      137  AACTGTGCCATCTGCAGGAACACACATATTGGATCTTTGCATAGAAATGTCAAGCTAACCCAG 196
QY      61  AlaSerAlaThrSerGluCysThrValAlaATpGlyValCysAsnHisAlaPheHis 80
Db      197  GCGTCGCTACTTCAGAAAGTGTACTGTCGATGGGAGTCTGTAAACCATCTTTTTCAC 256
QY      81  PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      257  TTCCACTGCATCTCTCGCTGGCTCAAAACACACAGAGTGTGTCCATTGGACACACAGAGAG 316
QY      101  TrpGluPheGlnLysTyrGlyHis 108
Db      317  TGGGAATTCAAAAGTATGGGCAC 340

RESULT 6
CQ711142
LOCUS      CQ711142
DEFINITION Sequence 56068 from Patent WO02070737.
ACCESSION CQ711142
VERSION    CQ711142.1 GI:42271999
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS   Liew, C.C., Marshall, W.E. and Zhang, H.
TITLE      Compositions and methods relating to osteoarthritis
            |||||||

```

Alignment Scores:									
Pred. No.:	1.41e-58	Length:	482						
Score:	616.00	Matches:	108						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	6	Gaps:	0						
US-09-541-462B-2 (1-108) x AX888031 (1-482)									
Qy	1	MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys	20						
Db	29	ATGGGGCAGCATGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAGAGAAG	88						
Qy	21	ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp	40						
Db	89	CGCTTTGAAGTGAAGTGAATGCAGTAGCCCTCTGGCCTGGGATATTGTGTTGAT	148						
Qy	41	AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln	60						
Db	149	AACTGTGCCATCTGCAGGAACCAATATGGATCTTTGCATAGAAATGTCAAAGCTAACAG	208						
Qy	61	AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis	80						
Db	209	GCCTCCGCTACTTCAGAGAGTGCTACTGTGCGCATGGGAGTCTGTAAACCATGCTTTTCAC	268						
Qy	81	PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu	100						
Db	269	TTCCACTGTCATCTCGCTGGCTCAAAACACGACAGGTGTGCCATTGGACAACAGAGAG	328						
Qy	101	TrpGluPheGlnLysTyrGlyHis	108						
Db	329	TGGGAATTCCAAAGATATGGGCAC	352						
RESULT 8									
LOCUS	BD027641	482 bp	DNA	linear	PAT 27-AUG-2002				
DEFINITION	Sequence tag and encoded human protein.								
ACCESSION	BD027641								
VERSION	BD027641.1 GI:22569383								
KEYWORDS	JP 2001269182-A/3887.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
TITLE	Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.								
JOURNAL	Patent: JP 2001269182-A 3887 02-OCT-2001;								
COMMENT	GENSET								
	OS	Homo sapiens (human)							
	PN	JP 2001269182-A/3887							
	PD	02-OCT-2001							
	PF	24-FEB-2000 JP 2000118773							
	PR	26-FEB-1999 US 60/122487							
	PI	JEAN BAPTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES							
	PI	JORDAN							
	PC	C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21,PC							
		C12N5/10,							
	PC	C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00,PC							
		G06F15/40							
CC	Location/Qualifiers								
FH	Key	29..352.							
FT	CDS	Location/Qualifiers							
		1..482							
FEATURES	source	/organism="Homo sapiens"							
		/mol_type="genomic DNA"							
		/db_xref="taxon:9606"							
ORIGIN	Alignment Scores:								

```

/db_xref="taxon:10095"

ORIGIN
Alignment Scores:
Pred. No.: 1.47e-58 Length: 504
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-541-462B-2 (1-108) x BD271522 (1-504)

Qy 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20
18 ATGCGCGCGCGATGGATGTGATATACCCCGCGCCACACAGCGCGCGCGCAAGAG 77
Qy 21 ArgPheGluValLysLysTyrAsnAlaValAlaLeuTyrPalaTyrAspIleValValAsp 40
78 CGCTTTGAAGTTAAAGATGGAATGCAATGCGATGCGCCCTCTGGCGCTGGGACATTTGTTGAT 137
Qy 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
138 AACTGTGCCATCTCAGAAACCAATATGATCTTTGTATCGAATGTCAGGCCAACCAAG 197
Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTyrGlyValCysAsnHisAlaPheHis 80
198 GCGTCAGCTACTTCCGAAGAGTGTCAGGTTCGATGGGAGGCTGCAACCATGCTTTTCAT 257
Qy 81 PheHisCysIleSerArgTyrLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
258 TTCACATGTCATCTCTCGATGGCTCAAAACAGGAGGAGGTGTGTCCGTTGGACAACAGAGAG 317
Qy 101 TrpGluPheGlnLysTyrGlyHis 108
318 TGGGAGTTCAGAAAGTATGGGCAT 341

RESULT 10
AF140599 504 bp mRNA linear ROD 11-MAY-1999
LOCUS Mus musculus ring-box protein 1 (Rbx1) mRNA, complete cds.
DEFINITION
ACCESSION AF140599
VERSION AF140599.1 GI:4769005
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 504)
Kamura,T., Koepf,D.M., Conrad,M.N., Skowrya,D., Moreland,R.J.,
Iliopoulos,O., Lane,W.S., Kaelin,W.G. Jr., Ellledge,S.J.,
Conaway,R.C., Harper,J.W., and Conaway,J.W.
RBx1, a component of the VHL tumor suppressor complex and SCF
ubiquitin ligase
Science 284 (5414), 657-661 (1999)
JOURNAL
MEDLINE 99234320
PUBMED 10213691
REFERENCE
2 (bases 1 to 504)
Kamura,T., Lane,W.S., Conaway,R.C. and Conaway,J.W.
Direct Submission
Submitted (05-APR-1999) Prog. Molec. Cell Biol., HHMI, Oklahoma
Med. Res. Fdn., 825 NE 13th St, Oklahoma City, OK 73104, USA
FEATURES
source
1..504
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..504
/gene="Rbx1"
18..344
/gene="Rbx1"
/notes="component of VHL tumor suppressor complex and SCF
ubiquitin ligase"

```

```
/codon_start=1
/product="ring-box protein 1"
/protein_id="RAD29716.1"
/db_xref="GI:4769306"
/translation="MAAAMDVDTFSGTNSGAGKGRFEVKKNVALMWDIVVDNCAL
CRNHIMDLICIEQANOASATSECTVANGVCHAFHFCISRWLKTRQVCPIDNREWE
FKYKGH"

ORIGIN

Alignment Scores:
Pred. No.: 1.47e-58 Length: 504
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-541-462B-2 (1-108) x AF140599 (1-504)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 18 ATGCGCGGGCGATGGATGGATATACCCAGCGGACCAACAGCGCGGCGGCAAGAAG 77
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
DB 78 CGCTTTGAAGTTAAAGTGAATGCAGTGGCCCTCTGGGCGCTGGGACATTGTGGTTGAT 137
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
DB 138 AACTGTGCCATCTGCAGGACCAACATATTATGGATCTTTGTATCGAATGTTCAGGCGCAACAG 197
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
DB 198 GCCTGAGCTACTTCCGAGAGTGATCGTTGCAATGGGAGTCTGCAACCATGCTTTTCAT 257
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
DB 258 TTCCACTGCATCTCTCGATGGCTCAAAACGAGGAGGTGTGTCGTTGGACACACAGAGAG 317
QY 101 TrpGluPheGlnLysTrpGlyHis 108
DB 318 TGGGAGTTCCAGAAGTATGGGCAT 341

RESULT 11
CQ729899 507 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 15833 from Patent WO02068579.
DEFINITION CQ729899
ACCESSION CQ729899
VERSION CQ729899.1 GI:42302243
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 15833 06-SEP-2002;
PE Corporation (NY) (US)

JOURNAL Location/Qualifiers
source 1..507
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1.48e-58 Length: 507
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-541-462B-2 (1-108) x CQ729899 (1-507)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 7 ATGCGCGGAGCGATGGATGTGATACCCGAGCGGACCAACAGCGGCGGCGCAAGAAG 66
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
DB 67 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 126
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
DB 127 AACTGTGCCATCTGCAGGACCAACATATTATGGATCTTTGCATAGATGTCAACCTAACACAG 186
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
DB 187 GCGTCCGCTACTTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTCAC 246
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
DB 247 TTCCACTGCATCTCTCGCTGGCTCAAAACGAGGAGTGTGTCATTTGGACACACAGAGAG 306
QY 101 TrpGluPheGlnLysTrpGlyHis 108
DB 307 TGGGAATTCCTCAAAAGTATGGGCAC 330

RESULT 12
BD271520 508 bp DNA linear PAT 17-JUL-2003
LOCUS VonHippel-Lindau tumor suppressor complex and novel component of
DEFINITION SCF ubiquitin ligase.
ACCESSION BD271520
VERSION BD271520.1 GI:33081288
KEYWORDS JP 2002541775-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Conaway,J.W., Conaway,R.C. and Kamura,T.
VonHippel-Lindau tumor suppressor complex and novel component of
SCF ubiquitin ligase
Patent: JP 2002541775-A 1 10-DEC-2002;
OKLAHOMA MEDICAL RESEARCH FOUNDATION

JOURNAL OS Homo sapiens (human)
COMMENT PN JP 2002541775-A/1
PD 10-DEC-2002
PF 25-FEB-2000 JP 2000601023
PR 26-FEB-1999 US 60/121787
PI JOAN W CONAWAY, RONALD C CONAWAY, TAKUMI KAMURA PC
C12N15/09,A61K38/00,A61K38/53,A61K45/00,A61P35/00,C07K14/47, PC
C12N1/15,
PC
C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12P21/02,G01N33/15,G01N33/ PC
50, G01N33/68
PC VonHippel-Lindau tumor suppressor complex and novel component
CC of SCF
CC ubiquitin ligase
CC Key Location/Qualifiers
FH Key 1..508
FT source /organism="Homo sapiens (human)".
FEATURES Location/Qualifiers
source 1..508
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
```



```

/translation="MAAAMDVTPSGTNSGAGKRFVKKNAVALMAWDIVVDNCA
CRNHMDLCIECQANQASATSECTVAMGVNCHAFPHFCISRLWKTQVCPLDNREWM
FKQYGH"

ORIGIN
Alignment Scores:
Pred. No.: 1.49e-58 Length: 508
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-541-462B-2 (1-108) x AF140598 (1-508)
QY 1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 7 ATGGCGGAGCGATGGATGGATACCCGAGCGGACCAACAGCGGCGGCGGCAAG 66
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspLleValValAsp 40
DB 67 CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTTGTTGAT 126
QY 41 AsnCySAlaLleCySArgAsnHisIleMetAspLeuCySileGluCySAlaAsnGln 60
DB 127 AACTGTGCCATCTCGAGAACCAATATGATCTTTGCATAGAAATGTCAGGTAACCCAG 186
QY 61 AlaSerAlaThrSerGluGluCySThrValAlaTrpGlyValCySAsnHisAlaPheHis 80
DB 187 GCGTCCGCTACTTCAGAAAGAGTGACTGTGCATGGGAGTCTGTACCATGCTTTTTCAC 246
QY 81 PheHisCySileSerArgTrpLeuLysThrArgGlnValCySProLeuAspAsnArgGlu 100
DB 247 TTCCACTGCATCTCTCGCTGCTCAAAACACGACGAGTGTCTCATTTGGACAACAGAGAG 306

RESULT 14
BC082183 514 bp mRNA linear VRT 10-SEP-2004
LOCUS Xenopus laevis cDNA clone MGC:98645 IMAGE:7200307, complete cds.
DEFINITION BC082183
ACCESSION BC082183
VERSION BC082183.1 GI:51980557
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 514)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
12454917
2 (bases 1 to 514)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahey,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villaion,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Pred. No.: 1.49e-58 Length: 508
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-541-462B-2 (1-108) x BD271520 (1-508)
QY 1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 7 ATGGCGGAGCGATGGATGGATACCCGAGCGGACCAACAGCGGCGGCGGCAAG 66
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspLleValValAsp 40
DB 67 CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTTGTTGAT 126
QY 41 AsnCySAlaLleCySArgAsnHisIleMetAspLeuCySileGluCySAlaAsnGln 60
DB 127 AACTGTGCCATCTCGAGAACCAATATGATCTTTGCATAGAAATGTCAGGTAACCCAG 186
QY 61 AlaSerAlaThrSerGluGluCySThrValAlaTrpGlyValCySAsnHisAlaPheHis 80
DB 187 GCGTCCGCTACTTCAGAAAGAGTGACTGTGCATGGGAGTCTGTACCATGCTTTTTCAC 246
QY 81 PheHisCySileSerArgTrpLeuLysThrArgGlnValCySProLeuAspAsnArgGlu 100
DB 247 TTCCACTGCATCTCTCGCTGCTCAAAACACGACGAGTGTGTCCATTTGGACAACAGAGAG 306

RESULT 13
AF140598 508 bp mRNA linear PRI 11-MAY-1999
LOCUS Homo sapiens ring-box protein 1 (RBX1) mRNA, complete cds.
DEFINITION AF140598
ACCESSION AF140598
VERSION AF140598.1 GI:4769003
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 508)
Kamura,T., Koepf,D.M., Conrad,M.N., Skowrya,D., Moreland,R.J.,
Iliopoulos,O., Lane,W.S., Kaelin,W.G. Jr., Ellledge,S.J.,
Conaway,R.C., Harper,J.W. and Conaway,J.W.
Rbx1, a component of the VHL tumor suppressor complex and SCF
ubiquitin ligase
Science 284 (5414), 657-661 (1999)
99234320
10213691
2 (bases 1 to 508)
Kamura,T., Lane,W.S., Conaway,R.C. and Conaway,J.W.
Direct Submission
Submitted (05-APR-1999) Prog. Molec. Cell Biol., HMI, Oklahoma
Med. Res. Fdn., 825 NE 13th St, Oklahoma City, OK 73104, USA
Location/Qualifiers
1..508
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..508
/gene="RBX1"
7..333
/feature="RBX1"
/note="ring finger-like protein; component of VHL tumor
suppressor complex and SCF ubiquitin ligase"
/codon_start=1
/product="ring-box protein 1"
/protein_id="AAD29715.1"
/db_xref="GI:4769004"

source
gene
CDS

```

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,  
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. A full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 3 (bases 1 to 514)  
 Klein,S. and Gerhard,D.S.  
 Direct Submission  
 Submitted (01-SEP-2004) National Institutes of Health, Xenopus Gene  
 Collection (XGC), National Institute of Child Health and Human  
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD  
 20892-7510, USA  
 NIH-MGC Project  
 Contact: XGC help desk  
 Email: cgabs-r@mail.nih.gov  
 Tissue Procurement: Dr. Igor Dawid  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 189 Row: f Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein.

## FEATURES

source  
 1. 514  
 Location/Qualifiers  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="MGC:98645 IMAGE:7200307"  
 /tissue\_type="Embryo, stage 17/19, Xenopus"  
 /clone\_lib="NICHG\_XGC\_Emb2"  
 /lab\_host="DH10B"  
 /notes="Vector: pCMV-SPORT6"  
 10..336  
 /codon\_start=1  
 /product="Unknown (protein for MGC:98645)"  
 /protein\_id="AAH82183.1"  
 /db\_xref="GI:51980558"  
 /translations="MAAAMDVDPSTGNSGAGKRPVKKNVALWAWDIVVDNCAI  
 CRNHJMDLCTECQANQASBSECTVANGVCNHFHFCISRWLKTQVCPLDREWE  
 FQXIGH"

## CDS

10..336  
 /product="Unknown (protein for MGC:98645)"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1 5e-58 Length: 514  
 Score: 616.00 Matches: 108  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BC082183 (1-514)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
 |||  
 DB 10 ATGGCGCGCGGATGGATGGATACCCCGCGGCGCACCAACAGCGCGCGGCAAGAAG 69  
 |||  
 QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
 |||  
 DB 70 CGCTTTGAAGTTAAAAGTGAATGCAGTGGCCCTCTGGGCCCTGGGACATTGTGGTTGAT 129  
 |||  
 QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60

Db 130 AAAGTGGCCATCTGCAGGAACACATATTATGATCTTTGTATCGAATGTGAGCCACAG 189  
 |||  
 QY 61 AlaSerAlaThrSerClnClnCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
 |||  
 Db 190 GCGTCAGCTACTTCGGAAGAGTGTACGGTGTGATGGGAGTCTGCAACCATGCTTTTCAT 249  
 |||  
 QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
 |||  
 Db 250 TTCCACTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCCTTGGACAACAGAGAG 309  
 |||  
 QY 101 TrpGluPheGlnLysTrpGlyHis 108  
 |||  
 Db 310 TGGGAGTTCAGGAAGTATGGCAT 333  
 |||  
 RESULT 15  
 CQ701366 523 bp DNA linear PAT 03-FEB-2004  
 LOCUS  
 DEFINITION Sequence 46292 from Patent WO02070737.  
 ACCESSION CQ701366  
 VERSION CQ701366.1 GI:42262133  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Liew,C.C., Marshall,W.E. and Zhang,H.  
 AUTHORS  
 TITLE Compositions and methods relating to osteoarthritis  
 JOURNAL Patent: WO 02070737-A 46292 12-SEP-2002;  
 Chondrogene Inc. (CA)  
 FEATURES  
 Location/Qualifiers  
 source  
 1..523  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1 53e-58 Length: 523  
 Score: 616.00 Matches: 108  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-541-462B-2 (1-108) x CQ701366 (1-523)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
 |||  
 Db 19 ATGGCGCGCGGATGGATGGATACCCCGAGCGGCGCACCAACAGCGCGCGGCAAGAAG 78  
 |||  
 QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
 |||  
 Db 79 CGCTTTGAAGTCAAAAAGTGAATGCAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT 138  
 |||  
 QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
 |||  
 Db 139 AACTGTGCCATCTGCAGGAACACACATTATGGATCTTTGGCATAGATGTCAGGTACACAG 198  
 |||  
 QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
 |||  
 Db 199 GCGTCCGCTACTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAAACCATGCTTTTTCAC 258  
 |||  
 QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
 |||  
 Db 259 TTCCACTGCATCTCTCGCTCGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAG 318  
 |||  
 QY 101 TrpGluPheGlnLysTrpGlyHis 108  
 |||  
 Db 319 TGGGAATTCCAAAAGTATGGGCAC 342  
 |||

Search completed: June 24, 2005, 09:58:04  
Job time : 3379 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 24, 2005, 06:33:57 ; Search time 442 Seconds  
(without alignment)  
1446.453 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVTPSGTNSGAGK.....KTRQVCPLDNREWFQKVGH 108

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_epool/US09541462/runat\_23062005\_122623\_9472/app\_query.fasta\_1.263  
-DB=N Geneseq\_16Dec04 -QFMT=fastcap -SUPFIX=ring -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09541462 @CGN 1.1 470 @runat\_23062005\_122623\_9472 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	616	100.0	327	3 AAA96882	Aaa96882 Nucleotid
2	616	100.0	482	3 AAC03896	Aac03896 Human sec
3	616	100.0	504	3 AAA74980	Aaa74980 DNA encod
4	616	100.0	506	12 ADQ87496	Adq87496 Human tum
5	616	100.0	506	12 ADQ87156	Adq87156 Human tum

6	616	100.0	506	13 ADQ84881	Adq84881 Human tum
7	616	100.0	508	3 AAA74978	Aaa74978 DNA encod
8	616	100.0	508	13 ACN40951	Acn40951 Tumour-as
9	611	99.2	476	9 ACH29979	Ach29979 Human tes
c 10	607	98.5	3726	13 ADS09913	Adso9913 Human the
c 11	607	98.5	4543	11 ACN90161	Acn90161 Breast ca
c 12	603	97.9	503	12 ADQ92179	Adq92179 Human aut
13	603	97.9	539	5 AAH97860	Aah97860 Murine 7-
14	603	97.9	586	5 AAH97862	Aah97862 Murine 7-
15	596	96.8	380	8 ABX39512	Abx39512 Bovine ES
c 16	586	95.1	4476	5 ABV25615	Abv25615 Human pro
17	576	93.5	485	3 AAA43288	Aaa43288 Xenopus s
c 18	573	93.0	5347	6 AAS94844	Aas94844 Human DNA
c 19	554	89.9	3484	12 ADF42703	Adf42703 Human Tes
c 20	554	89.9	3484	12 ADQ18564	Adq18564 Human sof
c 21	554	89.9	3484	13 ADQ80865	Adq80865 Human SPA
c 22	554	89.9	5111	10 ADB75558	Adb75558 Prostate
c 23	554	89.9	5371	12 ADQ23032	Adq23032 Human sof
24	551.5	89.5	557	5 AAH97861	Aah97861 Murine 7-
25	535	86.9	692	4 ABL22527	AbL22527 Drosophil
26	526.5	85.5	830	12 ADP22559	Adp22559 Sea-squir
27	523.5	85.0	527	10 ADK56883	Adk56883 Plant DNA
28	523.5	85.0	531	10 ADK56888	Adk56888 Plant DNA
29	523.5	85.0	586	10 ADK56887	Adk56887 Plant DNA
30	523.5	85.0	617	10 ADK56882	Adk56882 Plant DNA
31	523.5	85.0	691	10 ADK56880	Adk56880 Plant DNA
32	521	84.6	683	10 ADK56878	Adk56878 Plant DNA
33	520	84.4	673	10 ADK56879	Adk56879 Plant DNA
c 34	519.5	84.3	553	13 ACN55764	Acn55764 Cotton an
35	519.5	84.3	608	13 ADR59906	Adr59906 Cotton cd
36	515.5	83.7	390	6 ABQ85150	Abq85150 Arabidops
37	515.5	83.7	635	3 AAC39854	Aac39854 Arabidops
38	514.5	83.5	544	13 ADR59907	Adr59907 Cotton cd
39	513.5	83.4	731	10 ADK56881	Adk56881 Plant DNA
c 40	507.5	82.4	831	13 ADR61971	Adr61971 Cotton cd
c 41	507	82.3	1050	10 ADD17060	Add17060 DNA (SeqI
c 42	507	82.3	1050	10 ADD17933	Add17933 DNA (SeqI
c 43	507	82.3	1050	10 ADK55138	Adk55138 Plant DNA
c 44	507	82.3	1050	10 ADK56974	Adk56974 Plant DNA
45	501.5	81.4	3208	4 AAD12859	Aad12859 Human nov

ALIGNMENTS

RESULT 1  
AAA96882  
ID AAA96882 standard; DNA; 327 BP.

AC AAA96882;

XX  
DT 19-FEB-2001 (first entry)

XX Nucleotide sequence of human ring finger protein ROC1.

XX ROC1; ROC2; cullin; ring finger protein; APC11; APC complex; SCF pathway;  
XX cullin dependent ubiquitin ligase; CDK inhibitor Sic1 degradation;  
XX tumour; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 1..327  
FT /\*tag= a  
FT /product= "ROC1"

XX WO200058472-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US008592.

XX PR 31-MAR-1999; 99US-0127261P.

XX PR 22-NOV-1999; 99US-0166927P.

```

XX (UYN-) UNIV NORTH CAROLINA.
PA Xiong Y, Ohta T;
XX WPI; 2000-647235/62.
XX P-PSDB; AAB19160.
XX Novel nucleic acid encoding cullin regulating ring finger proteins,
PT termed as ROC proteins similar to anaphase-promoting complex 11, for
PT therapeutic and diagnostic use.
XX Claim 1; Fig 2A; 83pp; English.
XX The present sequence encodes a human ROC1 ring finger protein. The
CC specification also describes human ROC2. ROC1 and ROC2 are similar to
CC APC11, a subunit of the APC complex. The proteins stimulate cullin
CC dependent ubiquitin ligase activity. ROC1 functions in vivo as an
CC essential regulator of CDK inhibitor Sic1 degradation by the SCF
CC (undefined) pathway. ROC proteins are useful for screening bioactive
CC agents that interfere with the binding of ROC proteins with cullin
CC proteins. Pharmaceutical formulations comprising ROC proteins are useful
CC for diagnostic and therapeutic purposes, preferably for diagnosing and
CC treating tumours
XX
XX Sequence 327 BP; 85 A; 75 C; 94 G; 73 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,14e-66 Length: 327
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-541-462B-2 (1-108) x AAA96882 (1-327)
QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 1 ATGGCGCAGCGATGGATGGATATCCCGAGCGGCCACCACAGCGCGCGGCAAGAG 60
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
DB 61 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
DB 121 AACTGTGCCATCTCGCTGGCTCAAAACACGACAGGTGTCTTGCATAGAAATGTCAAGCTAACCCAG 180
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
DB 181 GCCTCCGCTACTTCAGNAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTCAC 240
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
DB 241 TTCACCTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTCTTGCATAGAAATGTCAAGCTAACCCAG 300
QY 101 TrpGluPheGlnLysTrpGlyHis 108
DB 301 TGGGAATTCCAAAGTATGGGCAC 324

RESULT 2
AAC03896
ID AAC03896 standard; cDNA; 482 BP.
XX
AC AAC03896;
XX
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 3894.
XX
XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

```

```

XX Homo sapiens.
OS EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GEST ) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX WPI; 2000-500381/45.
DR
XX P-PSDB; AAG03890.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX Claim 1; SEQ ID NO 3894; 71pp + Sequence Listing; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
XX
XX Sequence 482 BP; 116 A; 103 C; 121 G; 140 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 3,64e-66 Length: 482
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-541-462B-2 (1-108) x AAC03896 (1-482)
QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
DB 29 ATGGCGCAGCGATGGATGGATATCCCGAGCGGCCACCACAGCGCGCGGCAAGAG 88
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
DB 89 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 148
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
DB 149 AACTGTGCCATCTCGCTGGCTCAAAACACGACAGGTGTCTTGCATAGAAATGTCAAGCTAACCCAG 208
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
DB 209 GCCTCCGCTACTTCAGNAGAGTGTACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTCAC 268
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
DB 269 TTCACCTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTCTTGCATAGAAATGTCAAGCTAACCCAG 328
QY 101 TrpGluPheGlnLysTrpGlyHis 108

```

Db 329 TGGGAATTCCTCAAAAGTATGGGCAC 352

RESULT 3

AAA74980

ID AAA74980 standard; DNA; 504 BP.

XX

AC AAA74980;

XX

DT 02-JAN-2001 (first entry)

XX

DE DNA encoding a murine cullin-interacting RING-H2 finger protein (Rbx1).

XX

KW Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;

KW tumour suppressor; carcinoma; Ring box associated carcinoma;

KW von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;

KW cerebellar hemangioblastoma; hemangioma; retinal angiomas;

KW pheochromocytomas; ss.

XX

OS Mus sp.

XX

XX WO200050445-A1.

PN

XX

PD 31-AUG-2000.

XX

PF 25-FEB-2000; 2000WO-US004838.

XX

PR 26-FEB-1999; 99US-0121787P.

XX

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

PA

XX

PI Conaway JA, Conaway RC, Kamura T;

XX

DR WPI; 2000-572067/53.

XX

PT Cullin interacting RING-H2 finger protein, a component of von Hippel-

PT Lindau tumor suppressor complex and Skp1-Cdc53p-F-box protein (SCF)

PT ubiquitin ligase, useful for diagnosing and treating Ring box protein

PT associated carcinomas.

XX

PS Disclosure; Page 35; 37pp; English.

XX

CC The present sequence encodes a murine cullin-interacting RING-H2 finger

CC protein (Ring box protein), designated Rbx1. The human Rbx1 polypeptide

CC is a tumour suppressor. Human Rbx1 is useful for diagnosing a

CC predisposition of a patient to certain carcinomas. It is also useful for

CC treating Ring box protein associated carcinomas or augmenting

CC metabolically deficient system in animals. Human Rbx1 is also useful for

CC evaluating the effectiveness of a therapeutic treatment for Ring box

CC associated carcinomas. Human Rbx1 can be used to screen for agents which

CC augment or inhibit the activity of other cullin-containing ubiquitin

CC ligase and of the VHL (von Hippel-Lindau) complex controlling the

CC conjugation of ubiquitin or ubiquitin-like proteins to various sets of

CC target proteins. Carcinomas which may be treated include renal

CC carcinomas, cerebellar hemangioblastomas and hemangiomas, retinal

CC angiomas and pheochromocytomas

XX

SQ Sequence 504 BP; 117 A; 107 C; 137 G; 143 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,868-66	Length:	504
Score:	616.00	Matches:	108
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-541-462B-2 (1-108) x AAA74980 (1-504)

Qy 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20

Db 18 ATGGGGGGGGGATGGATGTGTGATACCCCGGCGCACACAGCGGGCGGGCAGAG 77

Qy 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40

Db 78 CGCTTTGAAGTTAAAAAGTGAATGCAGTGGCCCTCTGGGCTGGACATTGTGTTGAT 137

Qy 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60

Db 138 AACTGTGCCATCTGCAGGAACCAACCATTTATGATCTTTGTATCGAATGTTCAGGCCAACAG 197

Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80

Db 198 GCGTCAGCTACTTCCGAAGAGTGTACGGTTGTCATGGGGAGTCTGCAACCATGCTTTTCAT 257

Qy 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100

Db 258 TTCACATGCACTCTCTCGATGGCTCAAAACGAGGCGAGTGTGTCGTGGACACAGAGAG 317

Qy 101 TrpGluPheGlnLysTyrGlyHis 108

Db 318 TGGGAGTTCCAGAAAGTATGGGCAT 341

RESULT 4

ADQ87496

ID ADQ87496 standard; cDNA; 506 BP.

XX

AC ADQ87496;

XX

DT 07-OCT-2004 (first entry)

XX

DE Human tumour-associated antigenic target (TAT) cDNA sequence #4374.

XX

KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

KW cancer; cell proliferative disorder; gene; ss.

XX

OS Homo sapiens.

XX

XX WO2004060270-A2.

PN

XX

PD 22-JUN-2004.

XX

PF 15-OCT-2003; 2003WO-US029126.

XX

PR 18-OCT-2002; 2002US-0418988P.

XX

XX (GETH ) GENENTECH INC.

PA (WUTD/) WU T D.

PA (ZHOU/) ZHOU Y.

XX

PI Wu TD, Zhou Y;

XX

DR WPI; 2004-534300/51.

XX

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

PT preventing or treating cell proliferative disorders such as cancer.

XX

PS Claim 1; SEQ ID NO 4374; 5504pp; English.

XX

CC The present invention describes an isolated tumour-associated antigenic

CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide

CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of

CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%

CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -

CC (c). Also described: (1) an expression vector comprising the above

CC nucleic acid; (2) a host cell comprising the above expression vector; (3)

CC a process for producing a polypeptide; (4) an isolated polypeptide

CC comprising: (a) an amino acid sequence encoded by any of the above

CC nucleotide sequences; (b) an amino acid sequence encoded by the full-

CC length coding region of the above nucleotide sequences; or (c) a sequence

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

CC comprising the above polypeptide fused to a heterologous polypeptide; (6)

CC an isolated antibody that binds to the above polypeptide; (7) a process

CC for producing the antibody; (8) an isolated oligopeptide that binds to

CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)

CC binding organic molecule that binds to the above polypeptide; (10) a

CC composition of matter comprising the above (chimeric) polypeptide,

CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein  
CC described above; (15) methods of diagnosing the presence of a tumour in a  
CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.

XX Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,89e-66 Length: 506  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-541-462B-2 (1-108) x ADQ87496 (1-506)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 6 ATGGCGGCGAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAG 65  
QY 21 ArgPheGluValLysLysTTPAsnAlaValAlaLeuTTPAlaTTPAspLeuValAsp 40  
DB 66 CGTTTGAAGTGAAGAAGTGGATGAGTACCCCTCTGGGCCCTGGGATATGTGGTTGAT 125  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
DB 126 AACTGTGCCATCTGCAGAACACCATATTATGGATCTTTGCATAGATGTCAAGCTAACCCAG 185  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTTPGlyValCysAsnHisAlaPheHis 80  
DB 186 GCGTCCGCTACTTCAGAAGAGTGTACTGTCCATGGGAGTCTGTAAACCATGCTTTTCAC 245  
QY 81 PheHisCysIleSerArgTTPLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 246 TTCACATGCACTCTCTGCTGGCTCAAAACACGACAGTGTGTCCATTGGNCAACAGAGAG 305  
QY 101 TTPGluPheGlnLysTyrGlyHis 108  
DB 306 TGGGAATTCACAAAGTATGGGCAC 329

#### RESULT 5

ADQ87156  
ID ADQ87156 standard; cDNA; 506 BP.

XX AC ADQ87156;  
XX DT 07-OCT-2004 (first entry)  
XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #4032.  
XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
XX KW cancer; cell proliferative disorder; gene; ss.  
XX OS Homo sapiens.  
XX PN WO2004060270-A2.

XX 22-JUL-2004.  
PD 15-OCT-2003; 2003WO-US029126.  
XX PF 18-OCT-2002; 2002US-0418988P.  
XX PR (GETH ) GENENTECH INC.  
XX PA (WUTD/) WU T D.  
XX PA (ZHOU/) ZHOU Y.  
XX PI Wu TD, Zhou Y;  
XX XX WPI; 2004-534300/51.  
DR New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 4032; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to  
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein  
CC described above; (15) methods of diagnosing the presence of a tumour in a  
CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.

XX SQ Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,89e-66 Length: 506  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-541-462B-2 (1-108) x ADQ87156 (1-506)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20



Db 6 ATGGCGGCGAGCGATGGATGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAG 65  
Qy 21 ArgpHeGluVallylsystrpAenAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
Db 66 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 125  
Qy 41 AenCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
Db 126 AACTGTGCCATCTGCAGAGAACACATATATGGATCTTTGCATAGATGTCAAGCTAACCCAG 185  
Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
Db 186 GGGTGGCTACTTTCAGAGAGTGTACTGTGGATGGGAGTCTGTAAACATGCTTTTCAC 245  
Qy 81 PheHisCysIleSerArgTrpLeuYstrpLeuYstrpGlnValCysProLeuAspAsnArgGlu 100  
Db 246 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGTCACACAGAGAG 305  
Qy 101 TrpGluPheGlnYstrpGlyHis 108  
Db 306 TGGGAATTCCTCAAAAGTATGGGCAC 329

## RESULT 6

ADQ84881  
ID ADQ84881 standard; cDNA; 506 BP.

AC ADQ84881;

DT 07-OCT-2004 (first entry)

DE Human tumour-associated antigenic target (TAT) cDNA sequence #1695.

DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;  
KW cancer; cell proliferative disorder; gene; ss.

OS Homo sapiens.

XX WO2004060270-A2.

PN 22-JUL-2004.

XX 15-OCT-2003; 2003WO-US0297126.

PR 18-OCT-2002; 2002US-0418988P.

XX (GETH ) GENENTECH INC.

PA (WUTD/) WU T D.

PA (ZHOU/) ZHOU Y.

XX Wu TD, Zhou Y;

XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,  
PT preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 1695; 5504pp; English.

XX The present invention describes an isolated tumour-associated antigenic  
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide  
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of  
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%  
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-  
CC (c). Also described: (1) an expression vector comprising the above  
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)  
CC a process for producing a polypeptide; (4) an isolated polypeptide  
CC comprising: (a) an amino acid sequence encoded by any of the above  
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-  
CC length coding region of the above nucleotide sequences; or (c) a sequence  
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide  
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)  
CC an isolated antibody that binds to the above polypeptide; (7) a process  
CC for producing the antibody; (8) an isolated oligopeptide that binds to

CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)  
CC binding organic molecule that binds to the above polypeptide; (10) a  
CC composition of matter comprising the above (chimeric) polypeptide,  
CC antibody, oligopeptide or TAT binding organic molecule, in combination  
CC with a carrier; (11) an article of manufacture comprising a container and  
CC the composition of matter contained within the container; (12) methods of  
CC inhibiting the growth of a cell that expresses the above protein, where  
CC the growth of the cell is at least in part dependent upon a growth  
CC potentiating effect of the above protein; (13) a method of  
CC therapeutically treating a mammal having a cancerous tumour comprising  
CC cells that express the above protein; (14) a method of determining the  
CC presence of a protein in a sample suspected of containing the protein in a  
CC described above; (15) methods of diagnosing the presence of a tumour in a  
CC mammal; (16) a method for treating or preventing a cell proliferative  
CC disorder associated with increased expression or activity of the above  
CC protein; and (17) a method of binding an antibody, oligopeptide or  
CC organic molecule to a cell that expresses the protein described above.  
CC The TAT sequences have cytostatic activities, and can be used in gene  
CC therapy. The composition and methods are useful for diagnosing,  
CC preventing or treating cancer. The composition is also used for preparing  
CC a medicament for the therapeutic treatment or diagnostic detection of a  
CC cell proliferative disorder or cancer. The present sequence represents a  
CC human TAT cDNA sequence from the present invention.

SQ Sequence 506 BP; 126 A; 105 C; 124 G; 151 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 3,89e-66 Length: 506  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-541-462B-2 (1-108) x ADQ84881 (1-506)

Qy 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20

Db 6 ATGGCGGCGAGCGATGGATGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAG 65

Qy 21 ArgpHeGluVallylsystrpAenAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40

Db 66 CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 125

Qy 41 AenCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60

Db 126 AACTGTGCCATCTGCAGAGAACACATATATGGATCTTTGCATAGATGTCAAGCTAACCCAG 185

Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80

Db 186 GCGTCCGCTACTTCAAGAGAGTGTACTGTCCATGGGAGTCTGTAAACCATGCTTTTCAC 245

Qy 81 PheHisCysIleSerArgTrpLeuYstrpLeuYstrpGlnValCysProLeuAspAsnArgGlu 100

Db 246 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGTCACACAGAGAG 305

Qy 101 TrpGluPheGlnYstrpGlyHis 108

Db 306 TGGGAATTCCTCAAAAGTATGGGCAC 329

## RESULT 7

AAA74978

ID AAA74978 standard; DNA; 508 BP.

XX AAA74978;

AC AAA74978;

XX 02-JAN-2001 (first entry)

DE DNA encoding a human cullin-interacting RING-H2 finger protein (Rbx1).  
DE Cullin-interacting RING-H2 finger protein; Ring box protein; Rbx1;  
KW tumour suppressor; carcinoma; Ring box associated carcinoma;  
KW von Hippel-Lindau complex; ubiquitin conjugation; renal carcinoma;

KW cerebellar hemangioblastoma; hemangioma; retinal angioma;  
XX pheochromocytomas; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 7..333  
FT /tag= a  
FT /product= "cullin-interacting RING-H2 finger protein  
FT (Rbx1)"  
XX  
XX  
PN WO200050445-A1.  
XX  
PD 31-AUG-2000.  
XX  
PF 25-FEB-2000; 2000WO-US004838.  
XX  
PR 26-FEB-1999; 99US-0121787P.  
XX  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX  
PI Conaway JA, Conaway RC, Kamura T;  
XX  
DR WPI; 2000-572067/53.  
DR P-PSDB; AAB08813.  
XX  
XX Cullin interacting RING-H2 finger protein, a component of von Hippel-  
PT Lindau tumor suppressor complex and Skp1-cdc53p-F-box protein (SCF)  
PT ubiquitin ligase, useful for diagnosing and treating Ring box protein  
PT associated carcinomas.  
XX  
XX Claim 3; Page 35; 37pp; English.  
XX  
XX The present sequence encodes a human cullin-interacting RING-H2 finger  
CC protein (Ring box protein), designated Rbx1. The polypeptide is a tumour  
CC suppressor. Rbx1 is useful for diagnosing a predisposition of a patient  
CC to certain carcinomas. It is also useful for treating Ring box protein  
CC associated carcinomas or augmenting metabolically deficient system in  
CC animals. Rbx1 is also useful for evaluating the effectiveness of a  
CC therapeutic treatment for Ring box associated carcinomas. Rbx1 can be  
CC used to screen for agents which augment or inhibit the activity of other  
CC cullin-containing ubiquitin ligase and of the VHL (von Hippel- Lindau)  
CC complex controlling the conjugation of ubiquitin or ubiquitin- like  
CC proteins to various sets of target proteins. Carcinomas which may be  
CC treated include renal carcinomas, cerebellar hemangioblastomas and  
CC hemangiomas, retinal angioma and pheochromocytomas  
XX  
SQ Sequence 508 BP; 126 A; 106 C; 124 G; 152 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,91e-66 Length: 508  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-541-462B-2 (1-108) x AAA74978 (1-508)

Qy 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
Db 7 ATGCGCGCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGCAAGAG 66  
Qy 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
Db 67 CGCTTTGAAGTGAAGAAAGTGAATGCAGTATGACCCCTCGGGCCCTGGGATATTGGTGTAT 126  
Qy 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysGlnAlaAsnGln 60  
Db 127 AACTGTCCATCTGCAGGACCAACATTATGATCTTTGCATAGATGTCAGCTAACCA 186  
Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80

Db 187 GCGTCGCTACTTCAGAAAGAGGTACTGTGCGATCGGGAGTCTGTAAACATGCTTTTCAC 246  
Qy 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
Db 247 TTCCACTGCATCTCTCGCTGGCTCAAAACACACAGAGTGTGTCATTGGACACAGAGAG 306  
Qy 101 TrpGluPheGlnLysTyrglyHis 108  
Db 307 TGGGAATTCCTCAAAAGTATGGGCAC 330  
RESULT 8  
ACN40951  
ID ACN40951 standard; cDNA; 508 BP.  
XX  
AC ACN40951;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Tumour-associated antigenic target (TAT) cDNA DNA326980, SEQ ID NO:6072.  
XX  
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;  
KW tumour; diagnosis; cell proliferative disorder; breast cancer;  
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;  
KW central nervous system cancer; bladder cancer; pancreatic cancer;  
KW cervical cancer; melanoma; leukaemia; hybridisation probe;  
KW chromosome identification; chromosome mapping; gene mapping;  
KW gene therapy; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2004030615-A2.  
XX  
PD 15-APR-2004.  
XX  
XX 29-SEP-2003; 2003WO-US028547.  
XX  
PR 02-OCT-2002; 2002US-0414971P.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX Wu TD, Zhang Z, Zhou Y;  
PI WPI; 2004-347921/32.  
XX  
DR P-PSDB; ABM82365.  
XX  
XX New tumor-associated antigenic target polypeptides and nucleic acids,  
PT useful in preparing a medicament for treating or detecting a  
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or  
PT prostate cancer or tumor.  
XX  
XX Claim 1; SEQ ID NO 6072; 7273pp; English.  
XX  
XX The invention relates to human tumour-associated antigenic target (TAT)  
CC polypeptides, and their related nucleic acids. The TAT polypeptides are  
CC overexpressed in cancer tissues compared to normal tissues, and may thus  
CC serve as effective targets for the diagnosis and treatment of cancer in  
CC mammals. The invention also relates to nucleic acid and polypeptide  
CC sequences at least 80% identical to the TAT nucleic acids and  
CC polypeptides; expression vectors and host cells comprising a TAT nucleic  
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic  
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a  
CC TAT polypeptide; and methods and compositions for the treatment or  
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,  
CC antibodies, antagonists, binding molecules and compositions are useful  
CC for diagnosing or treating a cell proliferative disorder associated with  
CC increased TAT expression, particularly cancers such as breast cancer,  
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder  
CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX

Sequence 508 BP; 126 A; 106 C; 124 G; 152 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,91e-66 Length: 508  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-541-462B-2 (1-108) x ACN40951 (1-508)

QY 1 MetAlaAlaMerAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20  
DB 7 ATGGCGCGAGCATGATGTGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAG 66  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspLysValValAsp 40  
DB 67 CGCTTTGAAGTCAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGGTTGAT 126  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
DB 127 AACTGTGCCATCTGCAGGAACACATTTATGGATCTTTGCATAGAATGTCAAGCTAACCCAG 186  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 187 GCGTCGCTACTTCAAGAGAGTGTACTGTCGATGGGAGTCTGTAAACCATGCTTTTTCAC 246  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 247 TTCACCTGCATCTCTCGTGGCTCAAAACACACAGAGTGTGTCATTTGGACACAGAGAG 306  
QY 101 TrpGluPheGlnLysTrpGlyHis 108  
DB 307 TGGGAATTCAAAAGTATGGGCAC 330

RESULT 9

ACH29979  
ID ACH29979 standard; cDNA; 476 BP.  
XX AC ACH29979;  
XX DT 13-OCT-2003 (first entry)  
XX DE Human testis cDNA #365.  
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
XX KW genome mapping; biodiversity; genetic disorder.  
XX OS Homo sapiens.  
XX PN US2003073623-A1.  
XX PD 17-APR-2003.  
XX PF 30-JUL-2001; 2001US-00918995.  
XX PR 30-JUL-2001; 2001US-00918995.  
XX PA (DRMA/) DRMANAC R T.  
XX PA (LABA/) LABAT I.  
XX PA (STAC/) STACHE-CRAIN B.  
XX PA (DICK/) DICKSON M C.  
XX PA (JONE/) JONES L W.  
XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX DR WPI; 2003-615964/58.  
XX PT New polynucleotide sequences obtained from various cDNA libraries, useful  
XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
XX PT mapping, in the recombinant production of protein, or in generating  
XX PT antisense DNA or RNA.

XX

PS Claim 1; SEQ ID NO 17191; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20030073623

XX Sequence 476 BP; 119 A; 111 C; 124 G; 118 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.47e-65 Length: 476  
Score: 611.00 Matches: 107  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.19% Indels: 0  
DB: 9 Gaps: 0

US-09-541-462B-2 (1-108) x ACH29979 (1-476)

QY 2 AlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArg 21  
DB 75 GCGCAGCGCATGATGTGATACCCCGAGCGGCACCAACAGCGCGGCGCAAGAGCGC 134  
QY 22 PheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspLysValValAspAsn 41  
DB 135 TTTGAAGTCAAAAGTGAATGCAGTAGCCCTCTGGGCTGGGATATTGTGTTGATAC 194  
QY 42 CysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAla 61  
DB 195 TGTCCCATCTGCAGGAACCAACATTTATGGATCTTTGATAGTAATGTCAAGCTAACCGCG 254  
QY 62 SerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPhe 81  
DB 255 TCCGCTACTTCAGAGAGTGTACTGTGCTGGGAGTCTGTAAACCATGCTTTTCACCTTC 314  
QY 82 HisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgLysTrp 101  
DB 315 CACTGCATCTCTCGTGGCTCAAAACACAGCAGGTGTGTCCATTGGACACAGAGAGTGG 374  
QY 102 GluPheGlnLysTrpGlyHis 108  
DB 375 GAATTCAAAAGTATGGGCAC 395

RESULT 10

ADS09913/C  
ID ADS09913 standard; DNA; 3726 BP.  
XX AC ADS09913;  
XX DT 16-DEC-2004 (first entry)  
XX DE Human therapeutic DNA - SEQ ID 150.  
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;  
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;  
XX KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.  
XX OS Homo sapiens.  
XX

PN WO2004080148-A2.  
 XX 23-SEP-2004.  
 XX 30-SEP-2003; 2003WO-US030720.  
 XX 02-OCT-2002; 2002US-0416186P.  
 XX (NUVE-) NUVELO INC.  
 XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;  
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;  
 XX WPI: 2004-668857/65.  
 DR P-PSDB; ADS10597.  
 XX New polynucleotide, useful in preparing a composition for diagnosing or  
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,  
 PT aplastic anemia or cancer for promoting wound healing.  
 XX Claim 1; SEQ ID NO 150; 718pp; English.  
 XX The invention relates to a novel isolated polynucleotide and the encoded  
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,  
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may  
 CC be useful in preparing a composition for diagnosing or treating  
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell  
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting  
 CC wound healing. The molecules may also be utilised during gene therapy  
 CC procedures. The current sequence is that of a human therapeutic DNA of  
 CC the invention. The current sequence is not shown explicitly within the  
 CC specification but can be accessed from the WIPO web-site.  
 XX Sequence 3726 BP; 996 A; 955 C; 845 G; 930 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores: 7,64e-64 Length: 3726  
 Pred. No.: 607.00 Matches: 106  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 98.54% Indels: 0  
 Query Match: 13 Gaps: 0  
 DB: 13  
 US-09-541-462B-2 (1-108) x ADS09913 (1-3726)  
 QY 3 AlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPhe 22  
 DB 492 GCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCGCTTT 433  
 QY 23 GluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCys 42  
 DB 432 GAAGTGAAGAGTGAATGAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATTAAGTGT 373  
 QY 43 AlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSer 62  
 DB 372 GCCATCTGCAGGACCAACCATATTATGGATCTTTGGCATAGATGTCAAGTCAACAGGCGTCC 313  
 QY 63 AlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHis 82  
 DB 312 GCTACTTTCAGAGAGTGTACTGTGCATGGGAGTCTGTAAACCATGCTTTTCACTTCCAC 253  
 QY 83 CysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGlu 102  
 DB 252 TGCATCTCTGCTGGCTCAAAACACAGACAGTGTGTTCATTTGGACACAGAGAGTGGAA 193  
 QY 103 PheGlnLysTyrGlyHis 108  
 DB 192 TTCCAAAAGTATGGGCAC 175  
 RESULT 11  
 ACN90161/c  
 ID ACN90161 standard; DNA; 4543 BP.  
 XX

AC ACN90161;  
 XX 02-DEC-2004 (first entry)  
 XX Breast cancer related marker, seq id 11311.  
 DE Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.  
 XX Homo sapiens.  
 OS US2003099974-A1.  
 PN 29-MAY-2003.  
 XX 18-JUL-2002; 2002US-00198846.  
 XX 18-JUL-2001; 2001US-0306220P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Lillie J, Xu Y, Wang Y, Steinmann K;  
 PI WPI: 2003-787014/74.  
 XX Novel isolated polypeptide associated with breast cancer, useful for  
 PT detecting presence of polypeptide in sample, as a marker for breast  
 PT cancer.  
 XX Disclosure; SEQ ID NO 11311; 36pp; English.  
 PS The invention relates to an isolated polypeptide (I) associated with  
 CC breast cancer which is encoded by a nucleic acid molecule comprising a  
 CC nucleotide sequence (S1). Further disclosed is an antibody that binds to  
 CC the polypeptide of the invention. The activity of the polypeptide of the  
 CC invention may be described as cytostatic. The antibody is useful for  
 CC detecting the presence of (I) in a sample. Nucleic acid molecules of the  
 CC invention are useful in the detection of breast tumours. (I) is useful as  
 CC a marker for breast cancer and in breast cancer therapy. Sequences given  
 CC in records ACN78851-ACN92934 represent nucleic acid markers associated  
 CC with breast cancer. Note: The sequence listing does not form part of the  
 CC specification but may be obtained in electronic format from the USPTO web  
 CC site at seqdata.uspto.gov/sequence.html?docID=20030099974  
 XX Sequence 4543 BP; 1069 A; 1171 C; 1083 G; 1220 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores: 1e-63 Length: 4543  
 Pred. No.: 607.00 Matches: 106  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 98.54% Indels: 0  
 Query Match: 11 Gaps: 0  
 DB: 11  
 US-09-541-462B-2 (1-108) x ACN90161 (1-4543)  
 QY 3 AlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPhe 22  
 DB 1085 GCAGCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCGCTTT 1026  
 QY 23 GluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCys 42  
 DB 1025 GAAGTGAAGAGTGAATGAGTAGCCCTCTGGGCCTGGGATATTGTGGTTGATAACTGT 966  
 QY 43 AlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSer 62  
 DB 965 GCCATCTGCAGGACCAACCATATTATGGATCTTTGGCATAGATGTCAAGTCAACAGGCGTCC 906  
 QY 63 AlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHis 82  
 DB 905 GCTACTTTCAGAGAGTGTACTGTGCATGGGAGTCTGTAAACCATGCTTTTCACTTCCAC 846  
 QY 83 CysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGlu 102

Db 845 TGCATCTCTCGTGGCTCAAAACACGACGAGTGTGTCTCCATTGGACACAGAGAGTGGCAA 786

Qy 103 PheGlnLysTyrGlyHis 108  
 |||||  
 Db 785 TTCCAAAAGTAGTGGGCAC 768

RESULT 12

ADQ92179

ID ADQ92179 standard; DNA; 503 BP.

XX

AC ADQ92179;

XX

DT 07-OCT-2004 (first entry)

XX

DE Human autoantigen DNA fragment MPMGP800L05336.

XX

KW ds; autoantigen; antibody; hybridoma; biosensor chip;

KW extracorporeal differential diagnosis; autoimmune disease;

KW ribosomal protein; tubulin;

KW dolichyl-diphospho-oligosaccharide-glycosyl transferase;

KW multiple sclerosis; rheumatoid arthritis; epitope mapping;

KW affinity chromatography; electrophoresis; autoantibody adhesion;

KW RNA interference; RNAi.

XX

OS Homo sapiens.

XX

FN WO2004058972-A1.

XX

XX 15-JUL-2004.

XX

PF 23-DEC-2002; 2002WO-EP014731.

XX

PR 23-DEC-2002; 2002WO-EP014731.

XX

PA (THIE/) THIESEN H.

XX

PA (LORE/) LORENZ P.

XX

PI Thiesen H, Lorenz P;

XX

DR WPI; 2004-543459/52.

XX

XX New human DNA autoantigens, useful as assay, diagnostic, and prognostic

PT reagents and for treating autoimmune disease, also related expression

PT products and antibodies with similar uses.

XX

PS Claim 1; SEQ ID NO 160; 110pp; German.

XX

CC This invention describes novel human DNA autoantigens which are used to

CC produce recombinant expression vectors; prokaryotic or eukaryotic cells;

CC poly- or mono-clonal antibodies (Ab) specific; hybridomas that express

CC monoclonal Ab; biosensor chips having an addressable sequence pattern as

CC probes; medical or diagnostic instruments that include the biosensor; for

CC extracorporeal differential diagnosis of autoimmune diseases and

CC predisposition to them. The autoantigen polynucleotides encode ribosomal

CC proteins; tubulins; dolichyl-diphospho-oligosaccharide-glycosyl

CC transferases and proteins. The antibodies may be labelled conventionally

CC with radioisotopes, coloured or fluorescent groups, or a member of the

CC biotin/avidin pair, or colloidal gold. The autoantigens can be directed

CC against mitochondria, liver-kidney microsomes; histidyl-tRNA; nuclear

CC membrane; neutrophilin/cytoplasm; insect cells; epidermal intracellular

CC or basal membrane antigens; Golgi or cell nuclei, or associated with

CC multiple sclerosis or rheumatoid arthritis. They are useful for epitope

CC mapping; in affinity chromatography or electrophoresis; for diagnosis,

CC prognosis, control of treatment or therapeutic response of autoimmune

CC diseases, particularly in vitro differential diagnosis of autoimmune

CC diseases; to produce biosensor chips or for autoantibody adhesion.

CC Autoantigen DNA can be used for therapeutic RNA interference against

CC autoantibodies. Biochips that carry the new materials are useful in

CC medical and diagnostic instruments. ADQ92020-ADQ92280 represent human

CC autoantigens.

XX

Sequence 503 BP; 131 A; 102 C; 120 G; 150 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.53e-64 Length: 503

Score: 603.00 Matches: 105

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 97.89% Indels: 0

DB: 12 Gaps: 0

US-09-541-462B-2 (1-108) x ADQ92179 (1-503)

Qy 4 AlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPheGlu 23  
 |||||

Db 3 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAGCGCTTTCGA 62  
 |||||

Qy 24 VallyLysTyrAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAla 43  
 |||||

Db 63 GTGAAAAAGTGAATGCAGTAGCCCTCTGGGCTCTGGGATATTGTTGATAAATCTGTGCC 122  
 |||||

Qy 44 IleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAla 63  
 |||||

Db 123 ATCTGCAGGAACCAACATTATGGATCTTTGCAATGAATGTCAAGCTAACCCAGCGCTCCGCT 182  
 |||||

Qy 64 ThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPheHisCys 83  
 |||||

Db 183 ACTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAACCATGCTTTTCATCTCCACTGC 242  
 |||||

Qy 84 IleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGluPhe 103  
 |||||

Db 243 ATCTCTCGTGGCTCAAAACACGACAGTGTGTCTCATTTGGACAACAGAGAGTGGGAATTC 302  
 |||||

Qy 104 GlnLysTyrGlyHis 108  
 |||||

Db 303 CAAAAGTAGTGGGCAC 317  
 |||||

RESULT 13

AAH97860

ID AAH97860 standard; DNA; 539 BP.

XX

AC AAH97860;

XX

DT 10-OCT-2001 (first entry)

XX

DE Murine 7-transmembrane G-protein coupled receptor coding sequence #104.

XX

KW Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;

KW 7-transmembrane G-protein coupled protein receptor; ds.

XX

OS Mus sp.

XX

XX WO200160999-A1.

XX

PD 23-AUG-2001.

XX

PF 14-FEB-2001; 2001WO-US004700.

XX

PR 14-FEB-2000; 2000US-0182377P.

XX

PA (IMCL-) IMCLONE SYSTEMS INC.

XX

PA (UYPR-) UNIV PRINCETON.

XX

PI Lemischka IR, Witte L, Pereira DS;

XX

DR WPI; 2001-522596/57.

XX

XX DNA Sequences encoding 7-transmembrane G-protein coupled protein

PT receptors characteristic of hematopoietic stem cells, useful for treating

PT leukemia.

XX

PS Claim 1; Page 62; 176pp; English.

XX

CC The present invention relates to murine coding sequences for 7-

CC transmembrane G-protein coupled protein receptors (7TM-GPCRs). The

CC present sequence is one such murine 7TM-GPCR coding sequence. The present

CC sequence was derived from hematopoietic stem cells. The present sequence  
 CC and its corresponding protein are useful in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate 7TM-GPCR expression.  
 CC 7TM-GPCRs identify specific signalling molecules, to activate an effector  
 CC -signalling cascade that triggers an intracellular response and  
 CC eventually a biological effect  
 XX

SQ Sequence 539 BP; 128 A; 112 C; 150 G; 149 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.68e-64 Length: 539  
 Score: 603.00 Matches: 105  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.89% Indels: 0  
 DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x AAH97860 (1-539)

QY 4 AlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLysArgPheGlu 23  
 DB 1 GCATGGATGTGGATACCCCGCCGACCAACAGCGCGCGGCAAGAGCGCTTTGAA 60  
 QY 24 ValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAla 43  
 DB 61 GTTAAAGTGGATGAGTGGCCCTCTGGCCCTGGACATTGTGTTGATAACTGTGCC 120  
 QY 44 IleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAla 63  
 DB 121 ATCTCGAGGAACACATATTATGGATCTTTGTATCGAATGTGAGCCCAACAGCGCTCAGCT 180  
 QY 64 ThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisCys 83  
 DB 181 ACTTCGAGAGGTACGGTTGATGGGAGTCTGCAACATGCTTTTCATTTCCACTGC 240  
 QY 84 IleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGluPhe 103  
 DB 241 ATCTCTCGATGGTCAAAACGAGCGAGGTGTGTCCTGTGACAAACAGAGTGGGAGTTC 300  
 QY 104 GlnLysTyrGlyHis 108  
 DB 301 CAGAAGTATGGGCAT 315

RESULT 14

AAH97862  
 ID AAH97862 standard; DNA; 586 BP.

XX AC AAH97862;  
 XX AC  
 XX 10-OCT-2001 (first entry)  
 DE DE  
 DE Murine 7-transmembrane G-protein coupled receptor coding sequence #106.  
 XX XX  
 XX Murine; hematopoietic stem cell; signalling; vaccine; 7TM-GPCR;  
 XX 7-transmembrane G-protein coupled protein receptor; ds.  
 XX OS Mus sp.  
 XX WO200160999-A1.  
 XX 23-AUG-2001.  
 XX 14-FEB-2001; 2001WO-US004700.  
 XX 14-FEB-2000; 2000US-0182377P.  
 XX (IMCL-) IMCLONE SYSTEMS INC.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR, Witte L, Pereira DS;  
 XX WPI; 2001-522596/57.

PT DNA Sequences encoding 7-transmembrane G-protein coupled protein  
 PT receptors characteristic of hematopoietic stem cells, useful for treating  
 PT leukemia.  
 XX

PS Claim 1; Page 63; 176pp; English.

XX The present invention relates to murine coding sequences for 7-  
 CC transmembrane G-protein coupled protein receptors (7TM-GPCRs). The  
 CC present sequence is one such murine 7TM-GPCR coding sequence. The present  
 CC sequence was derived from hematopoietic stem cells. The present sequence  
 CC and its corresponding protein are useful in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate 7TM-GPCR expression.  
 CC 7TM-GPCRs identify specific signalling molecules, to activate an effector  
 CC -signalling cascade that triggers an intracellular response and  
 CC eventually a biological effect  
 XX

SQ Sequence 586 BP; 140 A; 122 C; 164 G; 159 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.89e-64 Length: 586  
 Score: 603.00 Matches: 105  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.89% Indels: 0  
 DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x AAH97862 (1-586)

QY 4 AlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArgPheGlu 23  
 DB 1 GCATGGATGTGGATACCCCGCCGACCAACAGCGCGCGGCAAGAGCGCTTTGAA 60  
 QY 24 ValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsnCysAla 43  
 DB 61 GTTAAAGTGGATGAGTGGCCCTCTGGCCCTGGACATTGTGTTGATAACTGTGCC 120  
 QY 44 IleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAla 63  
 DB 121 ATCTCGAGGAACACATATTATGGATCTTTGTATCGAATGTGAGCCCAACAGCGCTCAGCT 180  
 QY 64 ThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisCys 83  
 DB 181 ACTTCGAGAGGTACGGTTGATGGGAGTCTGCAACATGCTTTTCATTTCCACTGC 240  
 QY 84 IleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrpGluPhe 103  
 DB 241 ATCTCTCGATGGTCAAAACGAGCGAGGTGTGTCCTGTGACAAACAGAGTGGGAGTTC 300  
 QY 104 GlnLysTyrGlyHis 108  
 DB 301 CAGAAGTATGGGCAT 315

RESULT 15

ABX39512  
 ID ABX39512 standard; cDNA; 380 BP.

XX AC ABX39512;  
 XX AC  
 XX 20-FEB-2003 (first entry)  
 DE DE  
 DE Bovine EST associated with lactation/muscle/fat deposition #4677.  
 XX XX  
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 XX muscle deposition; fat deposition; genome mapping; gene identification;  
 XX gene analysis; cattle breeding.  
 XX OS Bos Taurus.  
 XX US2002137139-A1.  
 XX 26-SEP-2002.  
 XX 24-SEP-2001; 2001US-00960352.

XX 12-JAN-1999; 99US-0115707P.  
 PR 11-JAN-2000; 2000US-00480902.  
 XX  
 PA (BYAT/) BYATT J C.  
 PA (MATH/) MATHIALAGAN N.  
 PA (TAON/) TAO N.  
 PA (WARR/) WARREN W C.  
 XX  
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;  
 XX WPI; 2003-110599/10.  
 DR  
 XX  
 PT New nucleic acid associated with lactation, and muscle and fat  
 PT deposition, useful for genome mapping, gene identification and analysis,  
 PT cattle breeding, or for genetically improving cattle.  
 XX  
 PS Claim 2; SEQ ID NO 4677; 245pp; English.  
 XX  
 CC The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMPD), derived from  
 CC cattle, and the LMPD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,  
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are  
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic  
 CC acid linked to a promoter and a 3' non-translated sequence that  
 CC functions in the cell to cause termination of transcription and addition  
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 CC (2) determining a level or pattern of a molecule in a bovine cell or  
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a  
 CC complementary nucleic acid molecule obtained from the bovine cell or  
 CC tissue, where hybridisation between the marker nucleic acid and the  
 CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMPD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the  
 CC 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The  
 CC present sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139  
 XX  
 SQ Sequence 380 BP; 98 A; 87 C; 103 G; 92 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 7.57e-64 Length: 380  
 Score: 596.00 Matches: 103  
 Percent Similarity: 100.00% Conservative: 2  
 Best Local Similarity: 98.10% Mismatches: 0  
 Query Match: 96.75% Indels: 0  
 DB: 8 Gaps: 0

US-09-541-462B-2 (1-108) x ABX39512 (1-380)

Qy 4 AlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysIleValArgPheGlu 23  
 Db 4 GCGATGGATGTGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGCGCTTTGAA 63  
 Qy 24 VallysLysTyrPAsnAlaValAlaLeuTyrAlaTyrAspIleValValAspAsnCysAla 43  
 Db 64 GTGAAAAGTGGATGTCAGTAGCCCTCTGGGCTGGGATATTGTGGTTGATTAACGTGCC 123  
 Qy 44 IleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAlaSerAla 63  
 Db 124 ATCTGCAGGAACCAATTATGATCTTTGCATAGAAATGTCAAGCCAAACAGCGCTCCGCT 183  
 Qy 64 ThrSerGluGluCysThrValAlaTyrGlyValCysAsnHisAlaPheHisCys 83  
 Db 184 ACTTCTGAAGAGTGCAACCGTGGCGGCGTCTGTAAACCATGCTTTTTCACCTTCACCTGC 243

Qy 84 IleSerArgTyrPleuLysThrArgGlnValCysProLeuAspAsnArgGluTyrGluPhe 103  
 Db 244 ATCTCTCGCTGGCTCAAAACACACAGCAGGTGTCTCCGTTGGACAAACAGAGAGTGGGAATTC 303  
 Qy 104 GlnLysTyrGlyHis 108  
 Db 304 CAAAGGTATGGGCAC 318

Search completed: June 24, 2005, 09:01:45  
 Job time : 449 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 24, 2005, 09:58:09 ; Search time 487 Seconds  
(without alignment)  
1386.152 Million cell updates/sec

Title: US-09-541-462B-2  
Perfect score: 616  
Sequence: 1 MAAMDVDPGTSNGAGKK.....KTRQVCPLDNREWFQKXGH 108

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6067369 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09541462/runat\_23062005\_122626\_9578/app\_query.fasta\_1.263  
-DB=Published Applications NA -OPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09541462 @CIGN\_1.1.480 @runat\_23062005\_122626\_9578  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10J\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	616	100.0	433	17	US-10-242-535A-43377	Sequence 43377, A
2	616	100.0	433	18	US-10-085-783A-43377	Sequence 43377, A
3	616	100.0	453	17	US-10-242-535A-35025	Sequence 35025, A
4	616	100.0	453	18	US-10-085-783A-35025	Sequence 35025, A
5	616	100.0	467	17	US-10-242-535A-39933	Sequence 39933, A
6	616	100.0	467	18	US-10-085-783A-39933	Sequence 39933, A
7	616	100.0	471	17	US-10-242-535A-57254	Sequence 57254, A
8	616	100.0	471	18	US-10-085-783A-57254	Sequence 57254, A
9	616	100.0	472	17	US-10-242-535A-56068	Sequence 56068, A
10	616	100.0	472	18	US-10-085-783A-56068	Sequence 56068, A
11	616	100.0	504	21	US-10-913-917-5	Sequence 5, Appli
12	616	100.0	508	21	US-10-913-937-3	Sequence 3, Appli
13	616	100.0	523	17	US-10-242-535A-46292	Sequence 46292, A
14	616	100.0	523	18	US-10-085-783A-46292	Sequence 46292, A
15	611	99.2	476	10	US-09-918-995-17191	Sequence 17191, A
16	607	98.5	4543	14	US-10-198-846-11311	Sequence 11311, A
17	603	97.9	430	17	US-10-242-535A-54751	Sequence 54751, A
18	603	97.9	430	18	US-10-085-783A-54751	Sequence 54751, A
19	596	96.8	380	9	US-09-960-352-4677	Sequence 4677, Ap
20	586	95.1	4476	20	US-10-357-930-25604	Sequence 25604, A
21	573	93.0	5347	16	US-10-240-965-99	Sequence 99, Appl
22	554	89.9	3484	20	US-10-723-860-1383	Sequence 1383, Ap
23	554	89.9	5111	15	US-10-205-823-382	Sequence 382, App
24	554	89.9	5371	20	US-10-723-860-5852	Sequence 5852, Ap
25	545	88.5	468	17	US-10-242-535A-47656	Sequence 47656, A
26	545	88.5	468	18	US-10-085-783A-47656	Sequence 47656, A
27	524	85.1	840	19	US-10-767-701-12172	Sequence 12172, A
28	523.5	85.0	527	21	US-10-487-901-4266	Sequence 4266, Ap
29	523.5	85.0	531	21	US-10-487-901-4271	Sequence 4271, Ap
30	523.5	85.0	586	21	US-10-487-901-4270	Sequence 4270, Ap
31	523.5	85.0	617	21	US-10-487-901-4265	Sequence 4265, Ap
32	523.5	85.0	691	21	US-10-487-901-4263	Sequence 4263, Ap
33	523.5	85.0	1259	19	US-10-437-963-11395	Sequence 11395, A
34	522.5	84.8	619	18	US-10-425-114-25647	Sequence 25647, A
35	521	84.6	683	21	US-10-487-901-4261	Sequence 4261, Ap
36	520	84.4	673	21	US-10-487-901-4262	Sequence 4262, Ap
37	519.5	84.3	553	19	US-10-021-323-10545	Sequence 10545, A
38	519.5	84.3	608	19	US-10-767-795-687	Sequence 687, App
39	518.5	84.2	824	18	US-10-424-599-6074	Sequence 6074, Ap
40	515.5	83.7	390	9	US-09-770-791-20	Sequence 20, Appl
41	515.5	83.7	893	20	US-10-425-115-16831	Sequence 16831, A
42	514.5	83.5	544	19	US-10-767-795-688	Sequence 688, App
43	513.5	83.4	731	21	US-10-487-901-4264	Sequence 4264, Ap
44	512	83.1	711	19	US-10-767-701-12171	Sequence 12171, A
45	509.5	82.7	486	17	US-10-242-535A-52747	Sequence 52747, A

ALIGNMENTS

RESULT 1  
US-10-242-535A-43377  
; Sequence 43377, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: Chondrogene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017

```
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 43377
/ LENGTH: 433
/ TYPE: DNA
/ ORGANISM: Human
US-10-242-535A-43377

Alignment Scores:
Pred. No.: 1.68e-74 Length: 433
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-541-462B-2 (1-108) x US-10-242-535A-43377 (1-433)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
Db 20 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db 80 CGCTTTGAAGTCAAAAAGTGGAAATGCAGTAGCCCTCTGGGCCCTGGGATATTGGTTGAT 139
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db 140 AACTGTGCCATCTGCAGGAACACACATTATGGATCTTTGGCATAGAATGTCAAGCTAACCCAG 199
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
Db 200 GCGTCCGCTACTTCAGAGAGTGTACTGTCCGATGGGAGTCTGTAAACCATGCTTTTCAC 259
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db 260 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACACAGAGAG 319
QY 101 TrpGluPheGlnLysTrpGlyHis 108
Db 320 TGGGAATTCCTCAAAAGTATGGGCAC 343

RESULT 2
US-10-085-783A-43377
/ Sequence 43377, Application US/10085783A
/ Publication No. US20040037841A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2002
/ CURRENT APPLICATION NUMBER: US/10/085,783A
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 43377
/ LENGTH: 433
/ TYPE: DNA
/ ORGANISM: Human
US-10-085-783A-43377

Alignment Scores:
Pred. No.: 1.68e-74 Length: 433
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-541-462B-2 (1-108) x US-10-085-783A-43377 (1-433)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
Db 20 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 79
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db 80 CGCTTTGAAGTCAAAAAGTGGAAATGCAGTAGCCCTCTGGGCCCTGGGATATTGGTTGAT 139
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db 140 AACTGTGCCATCTGCAGGAACACACATTATGGATCTTTGGCATAGAATGTCAAGCTAACCCAG 199
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
Db 200 GCGTCCGCTACTTCAGAGAGTGTACTGTCCGATGGGAGTCTGTAAACCATGCTTTTCAC 259
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
Db 260 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACACAGAGAG 319
QY 101 TrpGluPheGlnLysTrpGlyHis 108
Db 320 TGGGAATTCCTCAAAAGTATGGGCAC 343

RESULT 3
US-10-242-535A-35025
/ Sequence 35025, Application US/10242535A
/ Publication No. US20040013663A1
/ GENERAL INFORMATION:
/ APPLICANT: ChondroGene Inc.
/ APPLICANT: Liew, C.C.
/ TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
/ FILE REFERENCE: 4231/2005
/ CURRENT APPLICATION NUMBER: US/10/242,535A
/ PRIOR FILING DATE: 2002-09-12
/ PRIOR APPLICATION NUMBER: US 10/085,783
/ PRIOR FILING DATE: 2002-02-28
/ PRIOR APPLICATION NUMBER: US 60/305,340
/ PRIOR FILING DATE: 2001-07-13
/ PRIOR APPLICATION NUMBER: US 60/275,017
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: US 60/271,955
/ PRIOR FILING DATE: 2001-02-28
/ NUMBER OF SEQ ID NOS: 58994
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 35025
/ LENGTH: 453
/ TYPE: DNA
/ ORGANISM: Human
US-10-242-535A-35025

Alignment Scores:
Pred. No.: 1.79e-74 Length: 453
Score: 616.00 Matches: 108
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-541-462B-2 (1-108) x US-10-242-535A-35025 (1-453)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20
Db 24 ATGGCGGCGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGCGGCAAGAAG 83
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
```

Db	84	CGCTTTGAAGTGAAAGTGGGAATGCAGTAGGCCCTCTGGGCTCGGATATTTGGTTGAT	143
Qy	41	AenCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAenGln	60
Db	144	AACTGTGCCATCTGCAGGAACACATATATGGATCTTTGCATAGATGTCAAGCTAACCAAG	203
Qy	61	AlaSerAlaThrSerGluGluCysThrValAlaTtpGlyValCysAsnHisAlaPheHis	80
Db	204	CGCTCCGCTACTTCAGAGAAGGTGACTGTGCGATGGGGAGTCTGTAAACCATGCTTTTCAAC	263
Qy	81	PheHisCysIleSerArgTtpLeuIlyThrArgGlnValCysProIeuAspAsnArgGlu	100
Db	264	TTCCACTGCATCTCTCGTGGCTCAAAACACACAGGAGTGTGCCATTGGACCAACAGAGAG	323
Qy	101	TtpGluPheGlnLysTyrGlyHis	108
Db	324	TGGGAATTCAAAAGTATGGGCAC	347

## RESULT 4

```

US-10-085-783A-35025
? Sequence 35025, Application US/10085783A
? Publication No. US20040037841A1
?
? GENERAL INFORMATION:
? APPLICANT: ChondroGene Inc.
? APPLICANT: Liew, C.C.
? TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
? FILE REFERENCE: 4231/2002
? CURRENT APPLICATION NUMBER: US/10/085,783A
? CURRENT FILING DATE: 2002-02-28
? PRIOR APPLICATION NUMBER: US 60/305,340
? PRIOR FILING DATE: 2001-07-13
? PRIOR APPLICATION NUMBER: US 60/275,017
? PRIOR FILING DATE: 2001-03-12
? PRIOR APPLICATION NUMBER: US 60/271,955
? PRIOR FILING DATE: 2001-02-28
? NUMBER OF SEQ ID NOS: 58994
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 35025
? LENGTH: 453
? TYPE: DNA
? ORGANISM: Human
US-10-085-783A-35025

```

Alignment Scores:	
Pred. No.:	1.79e-74
Score:	616.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	18
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
	Matches: 108
	Length: 453

US-09-541-462B-2 (1-108) x US-10-085-783A-35025 (1-453)

Qy	1	MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys	20
Db	24	ATGGCGGAGCGATGGATGTGGATACCCGACGCGCACCAACAGCGCGCGCGCAAGAAG	83
Qy	21	ArgPheGluValValLeuLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp	40
Db	84	CGCTTTGAAGTGGAAAGTGGGATGCGATAGCCCTCTGGCGCTGGGATATTGTGGTTGAT	143
Qy	41	AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln	60
Db	144	AACGTGCGCATCTGCAGGAACCACATATTGGATCTTTGCATAGAATGTCAAGCTAACCAAG	203
Qy	61	AlaSerAlaThrSerGluClyCysThrValAlaTrpGlyValCysAsnHisAlaPheHis	80
Db	204	CGGTCCGTACTTTCAGAGAGGTGACTGTCCGATGGGGAGTCTGTAAACCATGCTTTTTCAC	263
Qy	81	PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu	100
Db	264	TTCCACTGCATCTCTCGTGGCTCAAAACACAGCAGGTGTGTCTCAATTGCAACACAGAGAG	323

Qy	101	TTrGluPheGlnLysTyrGlyHis	108
Db	324	TGGGAATTCCAAAGATGGGCAC	347

RESULT 5

US-10-242-535A-39933

; Sequence 39933, Application US/10242535A

; Publication No. US20040013663A1

; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

Alignment Scores:		
Pred. No.:	1,868-74	Length: 467
Score:	616.00	Matches: 108
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	17	Gaps: 0

US-09-541-462B-2 (1-108) x US-10-242-535A-39933 (1-467)

Qy	1	MetalaalalaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys	20
Db	20	ATGCGCGCAGCGATGSGATGTGGATATCCCGAGCGGCACCAACAGCGCGCGCGCAAGAAG	79
Qy	21	ArgPheGluValLysLysTrpAsnAlaValalLeuTrpAlaTrpAspIleValValAsp	40
Db	80	CGCTTTGAGTGAANAAGTGAATGCGAGTAGCCCTCTGGGCTGGGATATTGTGTTGAT	139
Qy	41	AsnCysalatlleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnalalaAsnGln	60
Db	140	AACTGTGCCATCTGCAGGAACCAATTATGGATCTTTGCATAGAAATGTCAAGCTAAACAG	199
Qy	61	AlaSerAlaThrSerGluGluCysThrValalatrPglyValCysAsnHisalalaPheHis	80
Db	200	CGGTCCCGCTACTTCAGAAAGAGTACTGTCCGATGGGGAGTCTGTAAACCATGCTTTTTCAC	259
Qy	81	PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu	100
Db	260	TTCCACTGCGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACAACAGAGAG	319
Qy	101	TrpGluPheGlnLysTyrGlyHis	108
Db	320	TGGGAATTCCAAAAGTATGGGCAC	343

RESULT 6  
 US-10-085-783A-39933  
 ; Sequence 39933, Application US/10085783A  
 ; Publication No. US20040037841A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ChondroGene Inc.  
 ; APPLICANT: Liw. C.C.

; SEQ ID NO 57254					
; LENGTH: 471					
; TYPE: DNA					
; ORGANISM: Human					
US-10-242-535A-57254					
Alignment Scores:					
Pred. No.:	1.89e-74	Length:	471		
Score:	616.00	Matches:	108		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	17	Gaps:	0		
US-09-541-462B-2 (1-108) x US-10-242-535A-57254 (1-471)					
QY	1	MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys	20		
Db	17	ATGGCGGCAGCGATGTGATACCCGAGCGGCCAACACAGCGCGCGCAAGAAG	76		
QY	21	ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValasp	40		
Db	77	CGCTTTGAAGTGAATAAGTGCAATGCAGTAGCCCTCTGGGCCCTGGGATATTGTGGTTGAT	136		
QY	41	AnCysAlaIleCysArqAsnHisIleMetAspLeuCystileGluCySGlnAlaAsnGln	60		
Db	137	AACTGTGCATCTCGAGGAACCAATTATGATGATCTTTTCATAGAATGTCAAGCTAACCCAG	196		
QY	61	AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis	80		
Db	197	GGGTGCGTACTTTCAGAAAGTGTACTGCGCATGGGAGTCTGTAACCATGCTTTTCAC	256		
QY	81	PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspIleArgGlu	100		
Db	257	TTCACATGCATCTCTCGCTGGCTCAAACACAGCAGGTGTGTCCATTGGACACACAGAGAG	316		
QY	101	TrpGluPheGlnLysTyrrGlyHis	108		
Db	317	TGGGAATTCCAAAAAGTATGGGCAC	340		
RESULT 8					
US-10-085-783A-57254					
; Sequence 57254, Application US/10085783A					
; Publication No. US20040037841A1					
; GENERAL INFORMATION:					
; APPLICANT: ChondroGene Inc.					
; APPLICANT: Liew, C.C.					
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis					
; FILE REFERENCE: 4231/2002					
; CURRENT APPLICATION NUMBER: US/10/085,783A					
; CURRENT FILING DATE: 2002-02-28					
; PRIOR APPLICATION NUMBER: US 60/305,340					
; PRIOR FILING DATE: 2001-07-13					
; PRIOR APPLICATION NUMBER: US 60/275,017					
; PRIOR FILING DATE: 2001-03-12					
; PRIOR APPLICATION NUMBER: US 60/271,955					
; PRIOR FILING DATE: 2001-02-28					
; NUMBER OF SEQ ID NOS: 58994					
; SOFTWARE: PatentIn version 3.2					
; SEQ ID NO 57254					
; LENGTH: 471					
; TYPE: DNA					
; ORGANISM: Human					
US-10-085-783A-57254					
Alignment Scores:					
Pred. No.:	1.89e-74	Length:	471		
Score:	616.00	Matches:	108		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	18	Gaps:	0		

US-09-541-462B-2 (1-108) x US-10-085-783A-57254 (1-471)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 17 ATGGCGGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGGCAAGAG 76  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
DB 77 CGCTTTGAAGTGAAAGTGAATGCAGTAGCCCTCTGGGCGCTGGGATATTGTGTTGAT 136  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuValLysLysLysLysLysLys 60  
DB 137 AACTGTGCACTTCAGAGAGTGTACTGTGCAATGATGATGATGATGATGATGATGAT 196  
QY 61 AlaSerAlaThrSerGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 197 GCGTCCGCTACTTCAGAGAGTGTACTGTGCAATGATGATGATGATGATGATGATGAT 256  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 257 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACAGAGAG 316  
QY 101 TrpGluPheGlnLysTrpGlyHis 108  
DB 317 TGGGAATTCCTCAAAAGTATGGGCAC 340

## RESULT 9

US-10-242-535A-56068  
; Sequence 56068, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 56068  
; LENGTH: 472  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (437)..(437)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (455)..(455)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-56068

## Alignment Scores:

Pred. No.: 1.89e-74 Length: 472  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-09-541-462B-2 (1-108) x US-10-242-535A-56068 (1-472)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 17 ATGGCGGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGGCAAGAG 76

DB 20 ATGGCGGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGGCAAGAG 79  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
DB 80 CGCTTTGAAGTGAAAGTGAATGCAGTAGCCCTCTGGGCGCTGGGATATTGTGTTGAT 139  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuValLysLysLysLysLysLys 60  
DB 140 AACTGTGCACTTCAGAGAGTGTACTGTGCAATGATGATGATGATGATGATGATGAT 199  
QY 61 AlaSerAlaThrSerGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 200 GCGTCCGCTACTTCAGAGAGTGTACTGTGCAATGATGATGATGATGATGATGATGAT 259  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 260 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGCAACAGAGAG 319  
QY 101 TrpGluPheGlnLysTrpGlyHis 108  
DB 320 TGGGAATTCCTCAAAAGTATGGGCAC 343

## RESULT 10

US-10-085-783A-56068  
; Sequence 56068, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liaw, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 56068  
; LENGTH: 472  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (437)..(437)  
; OTHER INFORMATION: n is a, c, g, or t  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (455)..(455)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-085-783A-56068

## Alignment Scores:

Pred. No.: 1.89e-74 Length: 472  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-09-541-462B-2 (1-108) x US-10-085-783A-56068 (1-472)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 20 ATGGCGGAGCGATGGATGGATACCCCGAGCGGCACCAACAGCGCGGCAAGAG 79  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
DB 80 CGCTTTGAAGTGAAAGTGAATGCAGTAGCCCTCTGGGCGCTGGGATATTGTGTTGAT 139

QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysGlnAlaAsnGln 60  
DB 140 AACTGTGCCATCTGCAGGACCAACATTATGGATCTTTGCATAGATGTCAAGCTAACACG 199  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 200 GCCTCGCTACTTTCAGAAAGTGTACTGTGGTGGGAGTCTGTAAACCATGCTTTTTCAC 259  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 260 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACGAGTGTGTCCATTGGCAACACAGAG 319  
QY 101 TrpGluPheGlnLysTyrGlyHis 108  
DB 320 TGGGATTCAGAAAGTATGGGCAT 343

## RESULT 11

US-10-913-937-5  
; Sequence 5, Application US/10913937  
; Publication No. US20050019813A1  
; GENERAL INFORMATION:  
; APPLICANT: Conaway, Joan A.  
; APPLICANT: Conaway, Ronald C.  
; APPLICANT: Kamura, Takumi  
; APPLICANT: Oklahoma Medical Research Foundation  
; TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor  
; TITLE OF INVENTION: Complex and SCF Ubiquitin Ligase  
; FILE REFERENCE: 021044-004600US  
; CURRENT APPLICATION NUMBER: US/10/913,937  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: WO PCT/US00/04838  
; PRIOR APPLICATION NUMBER: US/09/914,324  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 504  
; TYPE: DNA  
; ORGANISM: Mus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (18)..(344)  
; OTHER INFORMATION: Rbx1  
US-10-913-937-5

Alignment Scores:  
Pred. No.: 2,07e-74 Length: 504  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0  
US-09-541-462B-2 (1-108) x US-10-913-937-5 (1-504)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 18 ATGGCGCGCGCATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAAG 77  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
DB 78 CGCTTTGAAGTTAAAGTGAATGCAGTGGCCCTCTGGGCCCTGGGACATTTGGTTGAT 137  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysGlnAlaAsnGln 60  
DB 138 AACTGTGCCATCTGCAGGACCAACATTATGGATCTTTGTATCGAATGTCAAGGCCAACACG 197  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 198 GCCTCAGCTACTTCCGAAGAGTGTACGGTTGCATGGGAGTCTGCAACCATGCTTTTCAT 257

QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 258 TTCCACTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCCGTTGGACAACACAGAG 317  
QY 101 TrpGluPheGlnLysTyrGlyHis 108  
DB 318 TGGGATTCAGAAAGTATGGGCAT 341

## RESULT 12

US-10-913-937-3  
; Sequence 3, Application US/10913937  
; Publication No. US20050019813A1  
; GENERAL INFORMATION:  
; APPLICANT: Conaway, Joan A.  
; APPLICANT: Conaway, Ronald C.  
; APPLICANT: Kamura, Takumi  
; APPLICANT: Oklahoma Medical Research Foundation  
; TITLE OF INVENTION: Novel Component of von Hippel-Lindau Tumor Suppressor  
; TITLE OF INVENTION: Complex and SCF Ubiquitin Ligase  
; FILE REFERENCE: 021044-004600US  
; CURRENT APPLICATION NUMBER: US/10/913,937  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR FILING DATE: 1999-02-26  
; PRIOR APPLICATION NUMBER: WO PCT/US00/04838  
; PRIOR APPLICATION NUMBER: US/09/914,324  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 508  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(333)  
; OTHER INFORMATION: Rbx1  
US-10-913-937-3

Alignment Scores:  
Pred. No.: 2,09e-74 Length: 508  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0  
US-09-541-462B-2 (1-108) x US-10-913-937-3 (1-508)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 7 ATGGCGCGCGCATGTGGATATCCCGAGCGGCACCAACAGCGCGCGGCAAG 66  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
DB 67 CGCTTTGAAGTTAAAGTGAATGCAGTGGCCCTCTGGGCCCTGGGATTTGGTTGAT 126  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysGlnAlaAsnGln 60  
DB 127 AACTGTGCCATCTGCAGGACCAACATTATGGATCTTTGCATAGAAATGTCAAGCTAACACG 186  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 187 GCCTCGCTACTTTCAGAAAGTGTACTGTGCATGGGAGTCTGTAAACCATGCTTTTCAC 246  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 247 TTCCACTGCATCTCTCGCTGGCTCAAAACACGACGAGTGTGTCCATTGGCAACACAGAG 306  
QY 101 TrpGluPheGlnLysTyrGlyHis 108

```

; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 46292
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Human
; US-10-085-783A-46292

Alignment Scores:
Pred. No.:      2,18e-74      Length:      523
Score:          616.00      Matches:    108
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:     0
DB:              18          Gaps:       0

US-09-541-462B-2 (1-108) x US-10-085-783A-46292 (1-523)

QY      1  MetAlaAlaAaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 30
DB      19  ATGGCGGCGAGCGATGTGATACCCGAGCGGCACCAACAGCGCGCGGCAAGAG 78

QY      21  ArgPheGluValLysLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
DB      79  CGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 138

QY      41  AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
DB      139  AACTGTGGCATCTCGAGGAACCAATTATGATCTTTGGCATAGATCTTGAAGCTAAGCTAAC 198

QY      61  AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
DB      199  GGGTCCGCTACTTCAGAAGAGTGTACTGCGCATGGGGAGTCTGTAACCATGCTTTTTCAC 258

QY      81  PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100
DB      259  TTCCACTGCATCTCTCGTGGCTCAAAACACGACAGGTGTGTCATTGGACAACAGAGAG 318

QY      101  TrpGluPheGlnLysTyTGlyHis 108
DB      319  TGGGAATTCCAAAAGATATGGGCAC 342

RESULT 15
US-09-918-995-17191
; Sequence 17191, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17191
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(476)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17191

```

Alignment Scores: 9.28e-74 Length: 476  
 Pred. No.: 611.00 Matches: 107  
 Score: 611.00 Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 99.19% Gaps: 0  
 DB: 10

US-09-541-462B-2 (1-108) x US-09-918-995-17191 (1-476)

QY	2	AlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLysArg	21
DB	75	GCAGCAGCGATGGATGTGGATACCCCGAGCGGCCAACACAGCGCGCGGGCAGAGAGCGC	134
QY	22	PheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAspAsn	41
DB	135	TTTGAAGTGAATAAGTGAATGCAGTAGCCCTCTGGCCCTGGGATATTGTGTTGATAAC	194
QY	42	CysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGlnAla	61
DB	195	TGTGCCATCTGCAGGAACACATTATGGATCTTTGCATAGATGTCAAGCTAACCGAGCG	254
QY	62	SerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHisPhe	81
DB	255	TCGCTACTTCAGAGAGTGTACTGTGCGATGGGAGTCTGTAACCATGCTTTTCACTTC	314
QY	82	HisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGluTrp	101
DB	315	CACATGATCTCTCGCTGGCTCAAAACACGACAGGTGTGTCCATTGGACACACAGAGATGG	374
QY	102	GluPheGlnLysTyrGlyHis	108
DB	375	GAATTCATAAAGTATGGGCAC	395

Search completed: June 24, 2005, 13:11:56  
 Job time : 489 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 24, 2005, 07:45:14 ; Search time 7486 Seconds

(without alignments)  
549.151 Million cell updates/sec

Title: US-09-541-462B-2

Perfect score: 616  
Sequence: 1 MAAMDVDPGTSNGAGKK.....KTRQVCPLDNREWFQKVGH 108

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh

-O=/cgn2\_1/USPTO\_spo1/US09541462/runat\_23062005\_122624\_9491/app.query.fasta\_1.263  
-DB=EST -OPWT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09541462@cgn 1 1 3437 @runat\_23062005\_122624\_9491 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	616	100.0	354	5	BY344429
2	616	100.0	358	2	BB871840
3	616	100.0	358	5	BY035215
4	616	100.0	360	5	BY045296
5	616	100.0	362	5	BY099338
6	616	100.0	362	5	BY136952
7	616	100.0	362	5	BY343565
8	616	100.0	364	5	BY036231
9	616	100.0	364	5	BY090529

10	616	100.0	364	5	BY101264
11	616	100.0	370	5	BY294825
12	616	100.0	371	5	BY320216
13	616	100.0	373	5	BY065587
14	616	100.0	374	5	BY294855
15	616	100.0	375	5	BY037694
16	616	100.0	375	5	BY304291
17	616	100.0	377	5	BY041215
18	616	100.0	380	5	BY091882
19	616	100.0	381	5	BY307337
20	616	100.0	383	5	BY062504
21	616	100.0	387	5	BY291952
22	616	100.0	397	2	BE849383
23	616	100.0	401	5	BY075709
24	616	100.0	405	4	BF956282
25	616	100.0	420	5	BY263438
26	616	100.0	424	2	BF682266
27	616	100.0	432	5	BY281504
28	616	100.0	436	5	BY271641
29	616	100.0	438	6	CD287433
30	616	100.0	445	2	BF231130
31	616	100.0	448	5	BY252236
32	616	100.0	454	5	BY241321
33	616	100.0	456	2	AW461667
34	616	100.0	458	5	BY254635
35	616	100.0	460	5	BY254633
36	616	100.0	462	7	CN305890
37	616	100.0	471	4	BG339057
38	616	100.0	473	5	BX283972
39	616	100.0	477	2	BF457442
40	616	100.0	478	6	CB857843
41	616	100.0	488	4	BG777485
42	616	100.0	497	7	CK393930
43	616	100.0	498	6	CB296618
44	616	100.0	500	7	CR518803
45	616	100.0	504	7	CK782343

#### ALIGNMENTS

RESULT 1

LOCUS BY344429

DEFINITION BY344429 RIKEN full-length enriched, whole joints Mus musculus CDNA

Clone L730004M21 5', mRNA sequence.

Accession BY344429

Version BY344429.1 GI:26573917

Keywords EST.

Source Mus musculus (house mouse)

Organism Mus musculus

Reference 1 (bases 1 to 354)

Authors

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Niki, K., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konegaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maitats, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Ring, B. Z., Ringwald, M., Sadelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wanhestedt, C., Wang, F., Watanabe, T., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,



Petrovsky, N., Pillai, R., Pontius, J.U., Oi, D., Ramechandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Takedale, R.D., Tomita, M., Verdardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

**TITLE**  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL MEDLINE**  
Nature 420, 563-573 (2002)

**PUBMED**  
22354683

**COMMENT**  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

**FEATURES**  
source  
1. .358  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="I530017M08"  
/sex="female"  
/tissue\_type="placenta"  
/dev\_stage="11 days pregnant adult"  
/clone\_lib="RIKEN full-length enriched, 11 days pregnant adult female placenta"

**ORIGIN**

Alignment Scores:			
Pred. No.:	9.06e-61	Length:	358
Score:	616.00	Matches:	108
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-541-462B-2 (1-108) x BY035215 (1-358)



## RESULT 5

BY099338  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 CONTACT: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Iehi, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN,  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

FEATURES  
 source

1. 362  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="K630127B15"  
 /clone\_lib="RIKEN full-length enriched, pooled tissues,  
 adult spleen, etc."  
 /note="(dev\_stage=adult,tissue\_type=spleen,sex=male),  
 (dev\_stage=adult,tissue\_type=kidney,sex=male),  
 (dev\_stage=adult,tissue\_type=testis,sex=male),  
 (dev\_stage=adult,tissue\_type=thymus,sex=male),  
 (dev\_stage=adult,tissue\_type=heart,sex=male),  
 (dev\_stage=adult,tissue\_type=colon,sex=male),  
 (dev\_stage=adult,tissue\_type=stomach,sex=male),  
 (dev\_stage=adult,tissue\_type=liver,sex=male),  
 (dev\_stage=13 days embryo,tissue\_type=whole  
 body,sex=mix),(dev\_stage=14 days embryo,tissue\_type=whole  
 body,sex=mix),(dev\_stage=16 days embryo,tissue\_type=whole  
 body,sex=mix),(dev\_stage=17 days embryo,tissue\_type=whole  
 body,sex=mix),(dev\_stage=15 days pregnant,  
 adult,tissue\_type=amion,sex=female),(dev\_stage=10 days  
 neonate,tissue\_type=brain,sex=mix),(dev\_stage=10 days  
 neonate,tissue\_type=thymus,sex=mix),(dev\_stage=10 days  
 neonate,tissue\_type=heart,sex=mix)"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,196-61 Length: 362  
 Score: 616.00 Matches: 108  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-541-462b-2 (1-108) x BY099338 (1-362)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
 DB 20 ATGGCGCGCGGATGGATGTGGATACCCCGCCACACAGCGCGCGCGCAAGAG 79  
 QY 21 ArgPheGluValLysLysLysAsnAlaValAlaLeuTTPAlaTTPAspIleValValAsp 40  
 DB 80 CGCTTTGAAGTTAAANAAGTGAATGCAGTGGCCCTCTGGGCTTGGACATTGTGGTTGAT 139  
 QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
 DB 140 AACTGTGCCATCTGCAGCAACACACATTATGATCTTTGTATGATGTTCAGGCCAACAG 199  
 QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTTPGlyValCysAsnHisAlaPheHis 80  
 DB 200 CGGTGAGTACTTCCGAAGAGGTGACGGTTGCATGGGAGTGTGCAACCATCTTTTCAT 259  
 QY 81 PheHisCysIleSerArgTTPLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
 DB 260 TTCACCTGCATCTCTCGATGGCTCAAAACAGGCGAGGTGTGTCCGTGGACACAGAGAG 319  
 QY 101 TTPGluPheGlnLysTyrGlyHis 108  
 DB 320 TGGGAGTTCAGAGAGTATGGGCAT 343

RESULT 6  
 BY136952  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 GI:26272503

BY136952  
 BY136952 RIKEN full-length enriched, 17.5 days embryo whole body  
 Mus musculus cDNA clone L930084016 5', mRNA sequence.  
 BY136952  
 BY136952.1 GI:26272503

KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 362)	
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, R., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierzki, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
MEDLINE	22354683	
PUBMED	12466851	
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Inotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. please visit our web site (http://genome.gsc.riken.go.jp) for further details. Location/Qualifiers 1. .362 /organism="Mus musculus"	
FEATURES		
source		

ORIGIN		
Alignment Scores:		
Pred. No.:	9,196-61	Length: 362
Score:	616.00	Matches: 108
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Indels:	0	
Query Match:	100.00%	Gaps: 0
DB:	5	
US-09-541-462B-2 (1-108) x BV136952 (1-362)		
QY	1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20	
Db	23 ATGGCGCGCGCGCATGGATGGATACCCCGCGGCCACACAGCGCGCGGCAAGAG 82	
QY	21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40	
Db	83 CGCTTTGAAGTTAAAAAGTGAATGCAGTGGCCCTCTGGGCTGGGACATTGTGTTGAT 142	
QY	41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60	
Db	143 AACTGTGCCATCTGCAGGAACACCATATATGGATCTTTGTATCGAATGTGCAGCCACACAG 202	
QY	61 AlaSerAlaThrSerGluGluCysThrValAlaLTrpGlyValCysAsnHisAlaPheHis 80	
Db	203 GGTTCAGCTACTTCCGAAGAGTGACGGTTGATGGGAGTCTGCAGCAACATGCTTTTCAT 262	
QY	81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100	
Db	263 TTCCACTGCATCTCTCGATGGTCTAAACAGGAGGAGTGTGTCCTGTTGGACACACAGAG 322	
QY	101 TrpGluPheGlnLysTrpGlyHis 108	
Db	323 TGGGAGTTCAGAGAGTATGGGCAT 346	

RESULT 7		
BY343565	362 bp	mRNA linear EST 12-DEC-2002
LOCUS	BY343565 RIKEN full-length enriched, whole joints	Mus musculus cDNA
DEFINITION	clone L230046F06 5', mRNA sequence.	
ACCESSION	BY343565	
VERSION	BY343565.1	GI:26573053
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 362)	
AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, R., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierzki, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perlea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, D. J., Reid, J., Ring, B. Z., Ringwald, M.,	

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

# JOURNAL

## MEDLINE

## PUBMED

## COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Tissues were provided by Vassilis Aidinis ( Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece ) whose assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

# FEATURES

## source

Location/Qualifiers  
1..362  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="L230046F06"  
/tissue\_type="whole joints"  
/clone\_lib="RIKEN full-length enriched, whole joints"

# ORIGIN

## Alignment Scores:

Pred. No.: 9,19e-61 Length: 362  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BY343565 (1-362)

# QY

1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20

Db 20 ATGGCGGGGGGATGGATGTGGATATCCCGGCGGCACACAGCGGGGGCGGCGGCGGAG 79  
Qy 21 ArgPheGluValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40  
Db 80 CGCTTTGAAGTTAAAAAGTGAATGACGTGGCCCTCTGGCGCTGGGACATTTGGTTGAT 139  
Qy 41 AsnCysAlaLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60  
Db 140 AACTGTGTCATCTCGAGAACACACATATGATCTTTGTATGCAATGTGAGCCCAACAG 199  
Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaLysLysLysLysLysLysLysLysLys 80  
Db 200 GCGTCAGCTACTTCCGAAGAGTGTACGGTTGATGCGGAGTCTGCAACATGCTTTTCAT 259  
Qy 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
Db 260 TTCCACTGCACTCTCTCGATGCTCAAAACGAGGAGGTGTGTCCGTTGCAACAACAGAG 319  
Qy 101 TrpGluPheGlnLysTyrGlyHis 108  
Db 320 TGGGAGTTTCCAGAAAGTATGGGCAT 343

# RESULT 8

## BY036231

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

BY036231 364 bp mRNA linear EST 06-DEC-2002  
BY036231 RIKEN full-length enriched, 14 days pregnant adult female  
amion Mus musculus cDNA clone 1530023F07 5', mRNA sequence.  
BY036231  
BY036231.1 GI:26141674  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 364)  
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,  
Chothia, C., Corbani, L.E., Cousins, S., Della, E., Dragani, T.A.,  
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gassi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
22354683  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

# TITLE

# JOURNAL

# MEDLINE

# PUBMED

# COMMENT



Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 Location/Qualifiers  
 1. 364  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="1530023f07"  
 /sex="female"  
 /tissue\_types="amion"  
 /dev\_stage="14 days pregnant adult"  
 /clone\_lib="RIKEN full-length enriched, 14 days pregnant  
 adult female amnion"

ORIGIN

Alignment Scores:  
 Pred. No.: 9,266-61 Length: 364  
 Score: 616.00 Matches: 108  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BY036231 (1-364)

Qy 1 MetAlaAlaMetAapValAapThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
 Db 22 ATGGCGCGCGATGGATGGATACCCCGCGGACCAACAGCGCGGCAAGAAG 81  
 Qy 21 ArgPheGluValLysLysTTPAsnAlaValAlaLeuTTPAlaTTPAspIleValValAsp 40  
 Db 82 CGCTTTGAAGTTAAAAGTGAATGCAGTGGCTCTGGGCCCTGGGACATGTGTGTAT 141  
 Qy 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
 Db 142 AACTGTGCCATCTCGAGGACCAACATATATGATCTTTGTATCGAATGTGCGGCCAACAC 201  
 Qy 61 AlaSerAlaThrSerGluGluCysThrValAlaTTPGlyValCysAsnHisAlaPheHis 80  
 Db 202 GCGTCAGCTACTTCGAGAGATGACGTTGCAATGGGAGTGTGCAACACCATGCTTTTCAT 261  
 Qy 81 PheHisCysIleSerArgTTPLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
 Db 262 TTCCACTGCATCTCTCGATGGCTCAAAACGAGGAGGTGTGTCTGCTTGGACACAGAG 321  
 Qy 101 TTPGluPheGlnLysTyrGlyHis 108  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 322 TGGGAGTTTCCAGAGTATGGCAT 345

RESULT 9  
 BY090529  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS

BY090529 364 bp mRNA linear EST 07-DEC-2002  
 musculus cDNA clone K630081M12 5', mRNA sequence.  
 BY090529  
 BY090529.1 GI:26199768  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Mus musculus  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 364)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V.,  
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
 Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
 Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,  
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
 Vervardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
 Wells, C., Wilming, L.G., Wynnshaw-Boris, A., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
 Sakazume, N., Sasaki, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
 Itoh, M., Kaga, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
 Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 12466851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome



Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

## FEATURES

source  
Location/Qualifiers  
1. .364  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K63008IM12"  
/tissue\_type="whole body"  
/dev\_stage="14 days embryo"  
/clone\_lib="RIKEN full-length enriched, 14 days embryo  
whole body"

## ORIGIN

## Alignment Scores:

Pred. No.: 9,266-61 Length: 364  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BY090529 (1-364)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
DB 23 ATGGCGCGCGCGATGGATGTGGATATCCCGCAGCAGCAGCGCGCGCGCGAAG 82  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspLeuValValAsp 40  
DB 83 CGCTTTGAAGTTAAAGTGGATGAGTGGCGCTCTGGGCGCTGGGACATGTGGTTGAT 142  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaHsngln 60  
DB 143 AACTGTGCCATCTGCAGGAACACATATTATGGATCTTTGTATCGATGTGCAGCCCAACAG 202  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
DB 203 CGGTGAGTACTTCCGAGAGAGTGTACGGTGTGATGGGAGTGTGCAACCATGCTTTTCAT 262  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
DB 263 TTCAGTGCATCTTCGATGGTTCGAACGAGGAGGTGTGTCCTTGGACACAGAGAG 322  
QY 101 TrpGluPheGlnLysTrpGlyHis 108  
DB 323 TGGGAGTTCAGAGATGATGGCAT 346

## RESULT 10

## BY101264

LOCUS BY101264 RIKEN full-length enriched, 15 days pregnant adult female  
DEFINITION amion Mus musculus cDNA clone K630136007 5', mRNA sequence.

ACCESSION BY101264

VERSION BY101264.1 GI:26211881

## KEYWORDS

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

## AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V.,  
Chochoia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
Gariboldi, M., Gassi, C., Godzik, A., Gough, J., Grimmond, S.,  
Gustincich, S., Hirokawa, N., Jackeon, I.J., Jarvis, E.D., Kanai, A.,  
Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,  
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,  
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehii, Y.,  
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683  
PUBMED  
12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
Hirozane, T., Imotani, K., Iehii, Y., Itoh, M., Kawai, J., Konno, H.,  
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
Hayashizaki, Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES  
source

Location/Qualifiers

1. .364  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K630136007"  
/sex="female"  
/tissue\_type="amion"  
/dev\_stage="15 days pregnant adult"  
/clone\_lib="RIKEN full-length enriched, 15 days pregnant  
adult female amion"

## ORIGIN

Alignment Scores:

Pred. No.:

9,266-61

Length:

364

Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BY101264 (1-364)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20  
Db 15 ATGGCGCGCGCGATGGATGGATACCCCGAGCGGCAACAGCGCGCGCAAGAAG 74  
QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
Db 75 CGCTTTGAAGTTAAAAGTGGATGCAGTGGCCCTCTGGCCCTGGGACATTGGTTGAT 134  
QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60  
Db 135 AACTGTGCCATCTGCAGGACCAACATATTGGATCTTTGTATCGATGTCCAGGCCAACAG 194  
QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80  
Db 195 GCCTCAGCTACTCCGAAGAGTGTACGTTGGATGGGAGTCTGCACCATGCTTTTCAT 254  
QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAsnArgGlu 100  
Db 255 TTCACCTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCGTTGGNACACAGAG 314  
QY 101 TrpGluPheGlnLysTrpGlyHis 108  
Db 315 TGGGAGTTCAGAAAGTATGGCCAT 338

## RESULT 11

LOCUS BY294825 370 bp mRNA linear EST 11-DEC-2002  
DEFINITION BY294825 RIKEN full-length enriched, visual cortex Mus musculus  
cDNA clone K530330G11 5', mRNA sequence.

ACCESSION BY294825

VERSION BY294825.1 GI:26485162

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 370)

Okazaki.Y., Furuno.M., Kasukawa.T., Adachi.J., Bono.H., Kondo.S.,  
Nikado.I., Osato.N., Saito.R., Suzuki.H., Yamana.K., I.,  
Kiyosawa.H., Yagi.K., Tomaru.Y., Hasegawa.Y., Nogami.A.,  
Schonbach.C., Gojobori.T., Baldarelli.R., Hill.D.P., Bult.C.,  
Hume.D.A., Quackenbush.J., Schriml.L.M., Kanapin.A., Matsuda.H.,  
Batalov.S., Beisel.K.W., Blake.J.A., Bradt.D., Brusic.V.,  
Chothia.C., Corbani.L.E., Cousins.S., Dalla.E., Dragani.T.A.,  
Fletcher.C.F., Forrest.A., Frazer.K.S., Gaasterland.T.,  
Gariboldi.M., Gissi.C., Godzik.A., Gough.J., Grimmond.S.,  
Gustincich.S., Hirokawa.N., Jackson.I.J., Jarvis.E.D., Kanai.A.,  
Kawaji.H., Kawasawa.Y., Kedzierski.R.M., King.B.L., Konagaya.A.,  
Kurochkin.I.V., Lee.Y., Lenhard.B., Lyons.P.A., Maglott.D.R.,  
Maltais.L., Marchionni.L., McKenzie.L., Miki.H., Negashima.T.,  
Numata.K., Okido.T., Pavan.W.J., Portea.G., Pesole.G.,  
Petrovsky.N., Pillai.R., Pontius.J.U., Qi.D., Ramachandran.S.,  
Ravasi.T., Reed.J.C., Reed.D.J., Reid.J., Ring.B.Z., Ringwald.M.,  
Sandelin.A., Schneider.C., Semple.C.A., Setou.M., Shimada.K.,  
Sultana.R., Takenaka.Y., Taylor.M.S., Teasdale.R.D., Tomita.M.,  
Verardo.R., Wagner.L., Wahlestedt.C., Wang.Y., Watanabe.Y.,  
Wells.C., Wilming.L.G., Wynshaw-Boris.A., Ynagisawa.M., Yang.I.,  
Yang.L., Yuan.Z., Zavolan.M., Zhu.Y., Zimmer.A., Carninci.P.,  
Hayatsu.N., Hirozane-Kishikawa.T., Konno.H., Nakamura.M.,  
Sakazume.N., Sato.K., Shiraki.T., Waki.K., Kawai.J., Aizawa.K.,  
Arakawa.T., Fukuda.S., Hara.A., Hachizume.W., Imotani.K., Ishii.Y.,  
Itoh.M., Kagawa.I., Miyazaki.A., Sakai.K., Sasaki.D., Shibata.K.,  
Shinagawa.A., Yasunishi.A., Yoshino.M., Waterston.R., Lander.E.S.,  
Rogers.J., Birney.B. and Hayashizaki.Y.

Analysis of the mouse transcriptome based on functional annotation

## TITLE

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
22354683  
12466851  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
Aizawa.K., Akimura.T., Arakawa.T., Carninci.P., Fukuda.S.,  
Hirozane.T., Imotani.K., Ishii.Y., Itoh.M., Kawai.J., Konno.H.,  
Miyazaki.A., Murata.M., Nakamura.M., Nomura.K., Numazaki.R.,  
Ohno.M., Sakai.K., Sakazume.N., Sasaki.D., Sato.K., Shibata.K.,  
Shiraki.T., Tagami.M., Waki.K., Watanishi.A., Muramatsu.M. and  
Hayashizaki.Y. Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Encyclopaedia Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Michela Pagiolini and Takao K. Hensch (  
Laboratory for Neuronal Circuit Development Brain Science Institute  
RIKEN 2-1 Hirotsawa, Wako-shi, Saitama 351-0198 Japan) whose  
assistance we gratefully acknowledge. Please visit our web site  
(http://genome.gsc.riken.go.jp) for further details.

## FEATURES

source

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="K530330G11"  
/tissue\_type="visual cortex"  
/clone\_lib="RIKEN full-length enriched, visual cortex"

## ORIGIN

Alignment Scores:  
Pred. No.: 9,46e-61 Length: 370  
Score: 616.00 Matches: 108  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-09-541-462B-2 (1-108) x BY294825 (1-370)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys 20

Db 30 ATGGCGCGCGCGATGGATGGATACCCCGAGCGGCAACAGCGCGCGCAAGAAG 89

QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40

Db 90 CGCTTTGAAGTTAAAAGTGGATGCAGTGGCCCTCTGGCCCTGGGACATTGGTTGAT 149

QY 41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60

Db 150 AACTGTGCCATCTGCAGGACCAACATATTGGATCTTTGTATCGATGTCCAGGCCAACAG 209

QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80

```

Db      210 GGGTACGACTCTCCAGAGGTACGGTGTGATGGGAGTGTGCACCATGCTTTTCAT 269
QY      81 PheHisCysIleSerArgTrpLeuYsThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      270 TTCCACTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCGTTGGACACAGAGAG 329
QY      101 TrpGluPheGlnYsTrpGlyHis 108
Db      330 TGGGAGTTCAGAAAGTATGGGCAT 353

RESULT 12
LOCUS   BY320216
DEFINITION
Musculus cDNA clone 1420104M17 5', mRNA sequence.
ACCESSION
VERSION  BY320216
KEYWORDS
SOURCE   Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 371)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawachi, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravelli, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Iehii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, N., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, N., Waki, K., Watahiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to

```

```

prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Takashi Ishikawa (Department of Surgery
2 Yokohama City University 3-9 Fukuura, Kanazawa-ku, Yokohama
236-0004 Japan) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
FEATURES             Location/Qualifiers
     source            1..371
     /organism="Mus musculus"
     /mol_type="mRNA"
     /strains="CS7BL/6J"
     /db_xref="taxon:10090"
     /clone="I420104M17"
     /cell_type="osteoclast-like cell"
     /clone_lib="RIKEN full-length enriched, osteoclast-like
     cell"
ORIGIN
Alignment Scores:      9.49e-61      Length:      371
Pred. No.:            616.00      Matches:      108
Score:                100.00%      Conservative: 0
Percent Similarity:   100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels:       0
Query Match:          5           Gaps:         0
DB:
US-09-541-462B-2 (1-108) x BY320216 (1-371)
QY      1 MetAlaAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyValaGlyLysLys 20
Db      30 ATGGCGGGCGGCGATGGATGTGGATATCCCCAGCGGCACCAACAGCGCGCGGCGAGAG 89
QY      21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40
Db      90 CGCTTTGAAGTTAAAAAGTGAATGCAGTGGCCCTCTGGCCCTGGGACATTGTGTTGAT 149
QY      41 AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln 60
Db      150 AACTGTGCCCATCTGCAGAACCAACATATTATGGATCTTTGTATCGAATGTGAGCCCAACCAG 209
QY      61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis 80
Db      210 GCGTCAGCTACTTCCGAAGAGTGTACGGTTGATGGGAGTCTGCAACCATGCTTTTCAT 269
QY      81 PheHisCysIleSerArgTrpLeuYsThrArgGlnValCysProLeuAspAsnArgGlu 100
Db      270 TTCCACTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCGTTGGACACAGAGAG 329
QY      101 TrpGluPheGlnYsTrpGlyHis 108
Db      330 TGGGAGTTCAGAAAGTATGGGCAT 353

RESULT 13
LOCUS   BY065587
DEFINITION
Musculus cDNA clone 1920043D08 5', mRNA sequence.
ACCESSION
VERSION  BY065587
KEYWORDS
SOURCE   Mus musculus (house mouse)

```



Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Michela Fagioli and Takao K. Hensch (

Laboratory for Neuronal Circuit Development Brain Science Institute

RIKEN 2-1 Hirosewa, Wako-shi, Saitama 351-0198 Japan) whose

assistance we gratefully acknowledge. Please visit our web site

(http://genome.gsc.riken.go.jp) for further details.

Location/Qualifiers

1. .374

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="K53033OK07"

/tissue type="visual cortex"

/clone\_lib="RIKEN full-length enriched, visual cortex"

QY 21 ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp 40  
 DB 90 CGCTTTGAAGTTAAAGTGAATGCAGTGGCCCTCTCGGCGCTGGGACATTGTGGTGTAT 149  
 QY 41 AenCySAIaIleCySAeAeAeHisIleMetAspLeuCySilleGluCySglnAlaAenGln 60  
 DB 150 AACTGTGCATCTGCAGGAACACCATTAATGATCTTTGTATCGAATGTGAGCCCAACGAG 209  
 QY 61 AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAenHisAlaPheHis 80  
 DB 210 GCGTCAGCTACTTCCGAAGAGGTGACGGTTCATGGCGAGTCTGCAACCATCTTTTCAT 269  
 QY 81 PheHisCysIleSerArgTrpLeuLysThrArgGlnValCysProLeuAspAenArgGlu 100  
 DB 270 TTCCACTGCATCTCTCGATGGCTCAAAACGAGCGAGGTGTGTCCGTGGACCAACGAGAG 329  
 QY 101 TrpGluPheGlnLysTyrGlyHis 108  
 DB 330 TGGGAGTTCCAGAGATGATGGGCAT 353

# RESULT 15

BY037694

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BY037694 375 bp mRNA linear EST 06-DEC-2002  
 BY037694 RIKEN full-length enriched, CRL-1722 L5178Y-R Mus musculus  
 CDNA clone I730002F15 5', mRNA sequence.

BY037694.1 GI:26143137

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 375)

Nikaido, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Hume, D.C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Flacher, C.F., Forrest, A., Frazer, K.S., Gassterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

## TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

US-09-541-462B-2 (1-108) x BY294855 (1-374)

QY 1 MetAlaAlaMetAspValAspThrProSerGlyThrAspSerGlyAlaGlyLysLys 20

DB 30 ATGGCGGCGCGGATGGATGGATATCCCCAGCGGCACCAACAGCGCGCGGCAAGAG 89

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES  
 source

1..375 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="DBA/2"  
 /db\_xref="taxon:10090"  
 /clone="I73002F15"  
 /cell\_line="CRL-1722 L5178Y-R"  
 /clone\_lib="RIKEN full-length enriched, CRL-1722 L5178Y-R"

ORIGIN

Alignment Scores:

Pred. No.:	9.63e-61	Length:	375
Score:	616.00	Matches:	108
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	5	Gaps:	0

US-09-541-462B-2 (1-108) x BY037694 (1-375)

Qy	1	MetAlaAlaMetAspValAspThrProSerGlyThrAsnSerGlyAlaGlyLysLys	20
Db	24	ATGGCGCGCGATGGATGTGGATACCCCGCGGCACGACGCGCGCGCAAGAAG	83
Qy	21	ArgPheGluValLysLysTrpAsnAlaValAlaLeuTrpAlaTrpAspIleValValAsp	40
Db	84	CGCTTTGAAGTTAAAGTGAATGCAGTGGCCCTCTGGGCGCTGGGACATTGTGGTTGAT	143
Qy	41	AsnCysAlaIleCysArgAsnHisIleMetAspLeuCysIleGluCysGlnAlaAsnGln	60
Db	144	AACTGTGCCATCTGCAGAACACCATTTATCGATCTTTGTATCGAATGTACGGCCACCCAG	203
Qy	61	AlaSerAlaThrSerGluGluCysThrValAlaTrpGlyValCysAsnHisAlaPheHis	80
Db	204	GCCTCAGCTACTTCCGAAGAGTGTACGTTTGCATGGGGAGTCTGCAACCATGCTTTTCAT	263
Qy	81	PheHisCysIleSerArgTrpLeuValThrArgGlnValCysProLeuAspAsnArgGlu	100
Db	264	TTCCACTGCACTCTCTCGATGGCTCAAAACGAGGCGAGGTGTCTCGTTGGACACAGAGAG	323
Qy	101	TrpGluPheGlnLysTyrGlyHis	108
Db	324	TGGAGTTCCAGAGTATGGCAT	347

Search completed: June 24, 2005, 12:02:58  
 Job time : 7494 secs